

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 18:34:08 ; Search time 2172.55 Seconds

(without alignments)  
10796.945 Million cell updates/sec

Title: US-10-082-502-16

Perfect score: 806

Sequence: 1 cgcgcccaagctcgagccaa.....aagagagagagagagacta 806

Scoring table: IDENTITY\_NUC

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database :

Listing first 45 summaries

```

1: gb.ba:*
2: gb.hg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*
29: em.vl:*
30: em.hg.hum:*
31: em.hg.in:*
32: em.hg.other:*
33: em.hg.mus:*
34: em.hg.pln:*
35: em.hg.rod:*
36: em.hg.mam:*
37: em.hg.vrt:*
38: em.sy:*
39: em.hggo.hum:*
40: em.hggo.mus:*
41: em.hggo.other:*
    
```

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	750.4	93.1	814	6	AX440456
2	750.4	93.1	814	9	AB015631
3	749.4	93.0	1210	6	AX464040
4	743.4	92.2	1080	6	AY032624
5	590.4	73.3	158198	10	BC008261
6	583.2	72.4	1017	10	AF186115
7	583.2	72.4	1017	10	AF186115
8	308.6	38.3	649	9	AF186113
9	288.4	35.8	824	9	BC001027
10	237.8	29.5	564	6	AX440375
11	206	25.6	155023	2	AC025574
12	206	25.6	182350	2	AC073896
13	206	25.6	184762	2	AC023500
14	177.2	22.0	207424	10	AC090489
15	174	21.6	101666	2	AC109891
16	147	18.2	147	6	AX333297
17	147	18.2	147	6	AX408688
18	126	15.6	126	6	AX333296
19	126	15.6	126	6	AX408687
20	112	13.9	69753	2	AC103156
21	94	11.7	95437	2	AC096788
22	84.4	10.5	2378	9	AK094445
23	56	6.9	622	5	AB055671
24	55.4	6.9	484	9	HSW803189
25	53.4	6.6	996	3	AY094698
26	51.4	6.4	55359	2	AC012854
27	51.4	6.4	172904	3	AC007414
28	51.4	6.4	273390	3	AE003831
29	48.6	6.0	3522	9	HSW804706
30	48.4	6.0	3833	6	AF399910
31	48.4	6.0	12393	6	AX346165
32	47.6	5.9	168994	9	AC097480
33	47.6	5.9	2640	9	AF184965
34	47.2	5.9	1534	9	HSW801279
35	47.2	5.9	1929	10	NMPD5931A
36	47.2	5.9	2446	9	HSW805337
37	47	5.8	168418	9	AC009403
38	46.8	5.8	94871	2	AC111353
39	46.8	5.8	129360	2	AC117079
40	46.4	5.8	1053	5	AF460207
41	46.4	5.8	3499	10	BC007476
42	46.4	5.8	183982	2	AC009825
43	46.2	5.7	63365	2	AC087437
44	46.2	5.7	94804	2	AL358935
45	46.2	5.7	144984	9	HS4333F14

ALIGNMENTS

RESULT 1  
LOCUS AX440456 814 bp DNA linear PAT 28-JUN-2002  
DEFINITION Sequence 309 from Patent WO0190154.  
ACCESSION AX440456  
VERSION AX440456.1 GI:21665266  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Xu,J., Mitcham,J.L., Harlocker,S.L., Dillon,D.C., Secret,H.,  
Lodes,M.J., Algate,P.A., Fling,S.P., Mannion,D., Benson,D.R. and  
Carter,D.

TITLE Compositions and methods for the therapy and diagnosis of ovarian cancer  
 JOURNAL Patent: WO 0190154-A 309 29-NOV-2001;  
 CORIXA CORPORATION (US)  
 FEATURES Location/Qualifiers  
 source 1. 814  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 210 a 194 c 241 g 169 t  
 ORIGIN

Query Match 93.1%; Score 750.4; DB 6; Length 814;  
 Best Local Similarity 99.0%; Pred. No. 4.9e-183;  
 Matches 754; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

OY 11 CTGGGGCCCAAGTGAAGATCCAGCGCTGCGCGCTGGGCGCAGCGGGCGCCCTGG 70
DB 52 CTGGGGCCCAAGTGAAGATCCAGCGCTGCGCGCTGGGCGCAGCGGGCGCCCTGG 111
OY 71 GACCAAGGTGAGCAACCCCGTTACCTTAARATGAAGAGCTGGGCTGGCTGGCCCTG 130
DB 112 GACGAGAGTGGAGCGACCCCATTAAGTGAAGATGAAGAGCTGGGCTGGCTGGCCCTG 171
OY 131 CTTCGGGGCCCTGCTGGAGACCCGCTGGGCTGGGAGAGAGCCAGAGATCCACTGTGA 190
DB 172 CTTCGGGGGGCCCTGCTGGAGACCCGCTGGGCTGGGAGAGAGCCAGAGATCCACTGTGA 231
OY 191 GCATGAGAGGCTCTGTGTGATGACATAGAAATGGAAATGGCCAGGTGAGCCCAAGAAG 250
DB 232 GCATGAGAGGCTCTGTGTGATGACATAGAAATGGAAATGGCCAGGTGAGCCCAAGAAG 291
OY 251 ACCATTGAGATGGGATCTTCCGGATCAATCCAGATGGCAGCAGTGGTGAAGTG 310
DB 292 ACCATTGAGATGGGATCTTCCGGATCAATCCAGATGGCAGCAGTGGTGAAGTG 351
OY 311 CCTTATGCCCCGTCAGAGGCCACCTGACAGAGCTGCTGAGAGATATGTACCGGATG 370
DB 352 CCTTATGCCCCGTCAGAGGCCACCTGACAGAGCTGCTGAGAGATATGTACCGGATG 411
OY 371 AAGGATATGGGGAACAGATTGATCTTCCACCATCCGAGAGACTAGCTAGTGTAGTG 430
DB 412 AAGGATATGGGGAACAGATTGATCTTCCACCATCCGAGAGACTAGCTAGTGTAGTG 471
OY 431 GCGCGGATGAGAGATCCAGTACCTGACCTACAGAGCATCCGATCCGATCAGATATT 490
DB 472 GCGCGGATGAGAGATCCAGTACCTGACCTACAGAGCATCCGATCCGATCAGATATT 531
OY 491 ACCGAGACCCCTCAAGTTGCTGTGAGACATTTGAGAGATTTAGAGATTAACATCAT 550
DB 532 ACCGAGACCCCTCAAGTTGCTGTGAGACATTTGAGAGATTTAGAGATTAACATCAT 591
OY 551 GAATCTTTTCCGAGAGGCTGACATATGTTAAAGACAACTTTCAGTAAAGCAACAT 610
DB 592 GAATCTTTTCCGAGAGGCTGACATATGTTAAAGACAACTTTCAGTAAAGCAACAT 651
OY 611 CTTTGTGACCAATCCCTGACATATGCGATGAGCTATGAACCACTGGAGCAGCCAC 670
DB 652 CTTTGTGACCAATCCCTGACATATGCGATGAGCTATGAACCACTGGAGCAGCCAC 711
OY 671 ACTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730
DB 712 ACTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771
OY 731 GTTTTACTGAATTAAGTGAAGAAATATGAACCAAGTA 772
DB 772 GTTTTACTGAATTAAGTGAAGAAATATGAACCAAGTA 813

```

RESULT 2  
 AB015631 814 bp mRNA linear PRI 26-JUL-2001  
 LOCUS AB015631  
 DEFINITION Homo sapiens mRNA for type II membrane protein, complete cds,  
 clone:HP10390.

ACCESSION AB015631  
 VERSION AB015631.1 GI:4586839  
 KEYWORDS type II membrane protein.  
 SOURCE Homo sapiens gastric adenocarcinoma CDNA to mRNA,  
 clone.lib:pkal-meta-1 clone:HP10390.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1  
 YOKOYAMA-Kobayashi M., Yamaguchi T., Sekine S. and Kato S.  
 Selection of cDNAs encoding putative type II membrane proteins on  
 the cell surface from a human full-length cDNA bank  
 Gene 228 (1-2), 161-167 (1999)  
 MEDLINE 99173880  
 REFERENCE 2 (bases 1 to 814)  
 Kato, S.  
 Direct Submission  
 Submitted (16-JUN-1998) Selski Kato, Research Institute of National  
 Rehabilitation Center for the Disabled, Department of  
 Rehabilitation Engineering; 4-1 Namiki, Tokorozawa, Saitama  
 359-8555, Japan (E-mail: selski@rehab.go.jp,  
 Tel:042-995-3100(ex.2568), Fax:042-995-3132)  
 FEATURES Location/Qualifiers  
 source 1. 814  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="HP10390"  
 /issue\_type="gastric adenocarcinoma"  
 /clone.lib="pkal-meta-1"  
 145..693  
 /codon\_start=1  
 /product="type II membrane protein"  
 /protein\_id="BAI76498.1"  
 /db\_xref="GI:4586840"  
 /translation="MKGWMLALLGALLGTAMARRSODLHCACRALVDELEWEIAD  
 VPKRTIOMSEFRINPOGOSVPEVPAARSBAHLTELEEFCDMKKEVQIDSTHR  
 KNYRVYGRNESESELDQGRIDSDISGTLKFACESTVEYEBELIEFBSREADNYK  
 DKLSKRDLCDHALHSHDEL"  
 BASE COUNT 210 a 194 c 241 g 169 t  
 ORIGIN

Query Match 93.1%; Score 750.4; DB 9; Length 814;  
 Best Local Similarity 99.0%; Pred. No. 4.9e-183;  
 Matches 754; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

OY 11 CTGGGGCCCAAGTGAAGATCCAGCGCTGCGCGCTGGGCGCAGCGGGCGCCCTGG 70
DB 52 CTGGGGCCCAAGTGAAGATCCAGCGCTGCGCGCTGGGCGCAGCGGGCGCCCTGG 111
OY 71 GACCAAGGTGAGCAACCCCGTTACCTTAARATGAAGAGCTGGGCTGGCTGGCCCTG 130
DB 112 GACGAGAGTGGAGCGACCCCATTAAGTGAAGATGAAGAGCTGGGCTGGCTGGCCCTG 171
OY 131 CTTCGGGGGGCCCTGCTGGAGACCCGCTGGGCTGGGAGAGAGCCAGAGATCCACTGTGA 190
DB 172 CTTCGGGGGGCCCTGCTGGAGACCCGCTGGGCTGGGAGAGAGCCAGAGATCCACTGTGA 231
OY 191 GCATGAGAGGCTCTGTGTGATGACATAGAAATGGAAATGGCCAGGTGAGCCCAAGAAG 250
DB 232 GCATGAGAGGCTCTGTGTGATGACATAGAAATGGAAATGGCCAGGTGAGCCCAAGAAG 291
OY 251 ACCATTGAGATGGGATCTTCCGGATCAATCCAGATGGCAGCAGTGGTGAAGTG 310
DB 292 ACCATTGAGATGGGATCTTCCGGATCAATCCAGATGGCAGCAGTGGTGAAGTG 351
OY 311 CCTTATGCCCCGTCAGAGGCCACCTGACAGAGCTGCTGAGAGATATGTACCGGATG 370
DB 352 CCTTATGCCCCGTCAGAGGCCACCTGACAGAGCTGCTGAGAGATATGTACCGGATG 411
OY 371 AAGGATATGGGGAACAGATTGATCTTCCACCATCCGAGAGACTAGCTAGTGTAGTG 430
DB 412 AAGGATATGGGGAACAGATTGATCTTCCACCATCCGAGAGACTAGCTAGTGTAGTG 471

```

```

QY 431 GGGCGAATGAGATTCAGTGAACCTGACCTACAGGATCCGATCGATGATATT 490
    |||
Db 472 GGGCGAATGAGATTCAGTGAACCTGACCTACAGGATCCGATCGATGATATT 531
QY 491 AGGGGACCCCTCAAGTTGGTGGAGAGCATTTGGAGGAATACAGAGATGAATCATT 550
    |||
Db 532 AGGGGACCCCTCAAGTTGGTGGAGAGCATTTGGAGGAATACAGAGATGAATCATT 591
QY 551 GAATTCCTTTCCGAGAGGCTGACAAATGTTAAAGACAACTTTCAGTAAGCAACAGAT 610
    |||
Db 592 GAATTCCTTTCCGAGAGGCTGACAAATGTTAAAGACAACTTTCAGTAAGCAACAGAT 651
QY 611 CTTTGTGACCATGCGCTGACATATGCGATGATGAGTATGAACCACTGGAGAGCCAC 670
    |||
Db 652 CTTTGTGACCATGCGCTGACATATGCGATGATGAGTATGAACCACTGGAGAGCCAC 711
QY 671 ACTGCGTTGATGATGATCAGCCCGAGAGGGGAAATGTTGGCAATGCTTTATATATAT 730
    |||
Db 712 ACTGCGTTGATGATGATCAGCCCGAGAGGGGAAATGTTGGCAATGCTTTATATATAT 771
QY 731 GTTTTACTGAAATTAATGAAAAAATATGAAACCAAAAGTA 772
    |||
Db 772 GTTTTACTGAAATTAATGAAAAAATATGAAACCAAAAGTA 813

RESULT 3
AX464040 1210 bp DNA linear PAT 16-JUL-2002
LOCUS Sequence 173 from Patent WO0140466.
ACCESSION AX464040
VERSION AX464040.1 GI:21899037
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E.,
Gao,W.Q., Gerlielsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,
Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,
Wood,W.L. and Zhang,Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
same
JOURNAL Patent: WO 0140466-A 173 07-JUN-2001;
Genentech Inc. (US)
FEATURES
source 1..1210
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 261 a 316 c 379 g 254 t
ORIGIN
Query Match 93.0%; Score 749.4; DB 6; Length 1210;
Best Local Similarity 98.9%; Pred. No. 9.6e-183;
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 11 CTTGGGCGCAAGTGAAGATCCAGCGCTGCGACGCGCTTGGGCGGCGGCGCTGG 70
    |||
Db 450 CTTGGGCGCAAGTGAAGATCCAGCGCTGCGACGCGCTTGGGCGGCGGCGCTGG 509
QY 71 GACCAAAAGGTGAGCAACCCCGTTACCTAAARATGAAGGGTGGGCTGGCGCCG 130
    |||
Db 510 GACCAAAAGGTGAGCAACCCCGTTACCTAAARATGAAGGGTGGGCTGGCGCCG 569
QY 131 CTTTGTGGGGGCGCTGCTGGGAACCCCGTGGCTCGAGAGAGCGAGATCTCACTGTGA 190
    |||
Db 570 CTTTGTGGGGGCGCTGCTGGGAACCCCGTGGCTCGAGAGAGCGAGATCTCACTGTGA 629
QY 191 GCATGAGAGGCTCTGCTGGATGAATGAGAAATTTGCCAGGTGAGACCCCAAGAAG 250
    |||
Db 630 GCATGAGAGGCTCTGCTGGATGAATGAGAAATTTGCCAGGTGAGACCCCAAGAAG 689
QY 251 ACCATTCAATGGAGATCTTCCGGAATCAATCGAATGGACGACGATGCTGGAGGTG 310

```

```

Db 690 ACCATTCAATGGAGATCTTCCGGAATCAATCGAATGGACGACGATGCTGGAGGTG 749
QY 311 CTTTATGCCCGCTCAGAGGCCACCTCAGAGAGCTCTGGAGAGATATGTGACCGGATG 370
    |||
Db 750 CTTTATGCCCGCTCAGAGGCCACCTCAGAGAGCTCTGGAGAGATATGTGACCGGATG 809
QY 371 AAGAGATATGGGAGACAGATGATGCTTCCAGCCATCGGAAGAACTACGTACTGTAGTG 430
    |||
Db 810 AAGAGATATGGGAGACAGATGATGCTTCCAGCCATCGGAAGAACTACGTACTGTAGTG 869
QY 431 GGGCGAATGAGATTCAGTGAACCTGACCTACAGAGCTCTGGAGAGATATGTGACCGGATG 490
    |||
Db 870 GGGCGAATGAGATTCAGTGAACCTGACCTACAGAGCTCTGGAGAGATATGTGACCGGATG 929
QY 491 AGGGGACCCCTCAAGTTGGTGGAGAGCATTTGGAGGAATACAGAGATGAATCATT 550
    |||
Db 930 AGGGGACCCCTCAAGTTGGTGGAGAGCATTTGGAGGAATACAGAGATGAATCATT 989
QY 551 GAATTCCTTTCCGAGAGGCTGACAAATGTTAAAGACAACTTTCAGTAAGCAACAGAT 610
    |||
Db 990 GAATTCCTTTCCGAGAGGCTGACAAATGTTAAAGACAACTTTCAGTAAGCAACAGAT 1049
QY 611 CTTTGTGACCATGCGCTGACATATGCGATGATGAGTATGAACCACTGGAGAGCCAC 670
    |||
Db 1050 CTTTGTGACCATGCGCTGACATATGCGATGATGAGTATGAACCACTGGAGAGCCAC 1109
QY 671 ACTGCGTTGATGATGATCAGCCCGAGAGGGGAAATGTTGGCAATGCTTTATATATAT 730
    |||
Db 1110 ACTGCGTTGATGATGATCAGCCCGAGAGGGGAAATGTTGGCAATGCTTTATATATAT 1169
QY 731 GTTTTACTGAAATTAATGAAAAAATATGAAACCAAAAGT 771
    |||
Db 1170 GTTTTACTGAAATTAATGAAAAAATATGAAACCAAAAGT 1210

```

```

RESULT 4
AY032624 1080 bp mRNA linear PRI 18-APR-2002
LOCUS Homo sapiens saposin-like protein mRNA, complete cds.
DEFINITION AY032624
ACCESSION AY032624
VERSION AY032624.1 GI:20196198
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Bornhauser,B.C., Olsson,P.-A. and Lindholm,D.
TITLE NSAP is a novel saposin-like protein that interacts with MIR and
stimulates neurite outgrowth
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1080)
AUTHORS Olsson,P.-A. and Lindholm,D.
TITLE Direct Submission
JOURNAL Submitted (13-APR-2001) Neuroscience, Uppsala University,
Husargatan 3, Uppsala 75123, Sweden
FEATURES
source 1..1080
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q15"
419..967
/note="NSAP: encodes a type 2 membrane signal and a
putative ER retention motif."
/product="saposin-like protein"
/protein_id="AAK3148.1"
/db_xref="GI:20196198"
/translation="MKGMWALLLGLALGTAAARSRSDILGCAALVDELEINQ
VDPKTIOMGSRINPDGOSVVEPIPAKRSAAHLTELEELICDRKKEGQIDPSTHR
KNYRVVGRNSESSELDIGIRIDISIGTLKFACESIVEYEDELIEFFSREADNVK

```

## CDS

BASE COUNT 238 a 269 c 343 g 230 t  
 ORIGIN  
 Query Match 92.2%; Score 743.4; DB 9; Length 1080;  
 Best Local Similarity 98.9%; Pred. No. 3.3e-181;  
 Matches 747; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

11 CTGGGGCCAAAGTAAAGTCACAGCGCTCTGCCACGCTTGGGCGACGGGGGCGCTGG 70  
 326 CTGGGGCCAAAGTAAAGTCACAGCGCTTCCAGCGCTTGGGCGACGGGGGCGCTGG 385  
 71 GACCAAGGTGAGCAGCAACCCCGTTACCTAAATGAAAGGCTGGGCTGGGCGCTGG 130  
 386 GAGCAGAGGTGAGCAGCAACCCCGTTACCTAAATGAAAGGCTGGGCTGGGCGCTGG 445  
 131 CTCTGGGGGCGCTCTGGGAGCCGCTGGGCTGGGAGGAGGAGGATCTCCACTGTGA 190  
 446 CTCTGGGGGCGCTCTGGGAGCCGCTGGGCTGGGAGGAGGAGGATCTCCACTGTGA 505  
 191 GCATGAGGGCTCTGGTGGATGAATGGAATGGAAATGGCCAGGTGAGCCCAAGAG 250  
 506 GCATGAGGGCTCTGGTGGATGAATGGAATGGAAATGGCCAGGTGAGCCCAAGAG 565  
 251 ACATTCAGATGGATCTTCCGGATCAATCCAGATGGCAGCCAGTCACTGGTGAAGTG 310  
 566 ACATTCAGATGGATCTTCCGGATCAATCCAGATGGCAGCCAGTCACTGGTGAAGTG 625  
 311 CCTATGCGCGCTCAAGAGCCGCTCAAGAGCTCTGGAGAGATATGTGACCGGATG 370  
 626 CCTATGCGCGCTCAAGAGCCGCTCAAGAGCTCTGGAGAGATATGTGACCGGATG 685  
 371 AAGGATATGGGAGACAGATGTATCTCCACCCATCGAAGAACTAGCTAGCTAGTG 430  
 686 AAGGATATGGGAGACAGATGTATCTCCACCCATCGAAGAACTAGCTAGCTAGTG 745  
 431 GCGCGGAATGGAGAAATCCAGTCACTGAGCACTCAAGGATCGGATCGGATCACTAT 490  
 746 GCGCGGAATGGAGAAATCCAGTCACTGAGCACTCAAGGATCGGATCGGATCACTAT 805  
 491 AGCGGACCCCTCAAGTTCGCTGAGACATGTGGAGAAATACGAGATGAACATCAT 550  
 806 AGCGGACCCCTCAAGTTCGCTGAGACATGTGGAGAAATACGAGATGAACATCAT 865  
 551 GAATCTTTTCCCGAGAGGCTGACAAATGTTAAAGACAACTTTGACGTAAGCAGACAT 610  
 866 GAATCTTTTCCCGAGAGGCTGACAAATGTTAAAGACAACTTTGACGTAAGCAGACAT 925  
 611 CTTTGGACCATGCGCTGACATATGCAATGAGCTATGAAACCACTGAGAGCCGAC 670  
 926 CTTTGGACCATGCGCTGACATATGCAATGAGCTATGAAACCACTGAGAGCCGAC 985  
 671 ACTGCGTATGATGATCAACCCCGAGAGGAGAAATGTTGCAATGCTTTTATATATAT 730  
 986 ACTGCGTATGATGATCAACCCCGAGAGGAGAAATGTTGCAATGCTTTTATATATAT 1045  
 731 GTTTTACTGAAATTAATGAAAAATATGAAAC 765  
 1046 GTTTTACTGAAATTAATGAAAAATATGAAAC 1080

RESULT 5  
 BC008261 798 bp mRNA linear ROD 07-AUG-2002  
 LOCUS BC008261  
 DEFINITION Mus musculus, transmembrane protein 4, clone MGC:6853  
 IMAGE:2650612, mRNA, complete cds.  
 ACCESSION BC008261  
 VERSION BC008261.1 GI:14198400  
 KEYWORDS MGC.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 798)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 Contact: MGC help desk  
 Email: [cga@bbs-remail.nih.gov](mailto:cga@bbs-remail.nih.gov)  
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)  
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

FEATURES  
 source  
 1..798  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /map="129.C57BL/6J.FVB/N"  
 /clone="MGC:6853 IMAGE:2650612"  
 /tissue="type="Mammary tumor, Brca1-/fl; MMTV-Cre model. 10 months old, gross tissue."  
 /clone\_id="NCI CGAP\_Mam3"  
 /lab\_host="DH10B"  
 /note="vector: pCMV-SPORT6"  
 115..663  
 /codon\_start=1  
 /product="transmembrane protein 4"  
 /protein\_id="AA08261.1"  
 /db\_xref="GI:14198401"  
 /db\_xref="locusID:56530"  
 /translation="MKGWMLALLGLVLTGTAARRSODLHCGACRALVDELEWELAR VDPKRTIOMGSEFINPDGOSVVEVPARSAHLTELEVCDBMKEGEIDPSTHR KNYVAVSRNCESSLIDGIRIDISIDSLTGFACESIVEBEDELIEFFSREADNVK DKLSKRTDLCDDHALHSHDEL"

BASE COUNT 227 a 182 c 228 g 161 t  
 ORIGIN  
 Query Match 73.3%; Score 590.4; DB 10; Length 798;  
 Best Local Similarity 87.0%; Pred. No. 1.1e-141;  
 Matches 672; Conservative 1; Mismatches 92; Indels 7; Gaps 2;

18 CAAAGTGAAGTCACAGCGCTCTGCCACGCTTGGGCGACGGGGGCGCCCTGGACCAA 77  
 29 CAAAGTGAAGTCACAGCGCTCTGCCACGCTTGGGCGACGGGGGCGCCCTGGACCAA 88  
 78 GGTGAGCAGCAACCCCGTTACCTAAATGAAAGGCTGGGCTGGGCTGGGCTGGGCTGG 137  
 89 GGTGAGCAGCAACCCCGTTACCTAAATGAAAGGCTGGGCTGGGCTGGGCTGGGCTGG 148  
 138 GGGCCCTGCTGGGAACCCGCTGGGCTGGGAGAGCCAGGATCTCCACTGTGAGCATGCA 197  
 149 GGGCTCTGCTGGGAACCCGCTGGGCTGGGAGAGCCAGGATCTCCACTGTGAGCATGCA 208  
 198 GGGCTCTGCTGGATGAATGAGATGGGAATGGCCAGGTGAGCAACCCCAAGAGACATTC 257  
 209 GGGCTCTGCTGGATGAATGAGATGGGAATGGCCAGGTGAGCAACCCCAAGAGACATTC 268  
 258 AGATGGATCTTCCGATCAATCCAGATGGAGGAGCAGTGGTGGAGGAGGCTTATG 317



```

Db 269 AGATGGGATCCTTCGCAATCATCCAGATGGCAGCCAGCATGTTGAGAGTACCTTANG 328
Oy 318 CCCGCTCAGAGGCCACCTCAGACAGCTCCTGAGAGATATGTGACCCGATGAGAGACT 377
Db 329 CCCGCTCAGAGGCCACCTCAGACAGCTCCTGAGAGATATGTGACCCGATGAGAGACT 388
Oy 378 ATGGGGAACAGATTTGATCTTCACCCATCGCAAGAACTACGATGTAAGTGGCCGGA 437
Db 389 ACGGGGAACAGATTTGATCTTCACCCATCGCAAGAACTACGATGTAAGTGGCCGGA 448
Oy 438 ATGGGAATTCAGTGAATGAGACCTTACCAAGCATCCGAATTCAGATATATAGGCGCA 497
Db 449 ATGGGAATTCAGTGAATGAGACCTTACCAAGCATCCGAATTCAGATATATAGGCGCA 508
Oy 498 CCGTCAAGTTTGGCTGAGAGCATTTGTGAGAGAAATACGAGATGAATCTTGAATTC 557
Db 509 CCGTCAAGTTTGGCTGAGAGCATTTGTGAGAGAAATACGAGATGAATCTTGAATTC 568
Oy 558 TTTCCGGAAGGCTGACATGTTTAAAGCAAACTTTGCGATAGGCAACAGATCTTTG 617
Db 569 TCTCCAGAGAGGCTGACATGTTTAAAGCAAACTTTGCGATAGGCAACAGATCTTGT 628
Oy 618 ACCATCCCTGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 672
Db 629 ACCATCCCTGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 688
Oy 673 TGGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 732
Db 689 AAACGTGATGAGAACCCCGCAGAGGGAAGATGGCAGATGTCCTTT--TATATTACGT 746
Oy 733 TTTTACTGAAATTAATGAAAAATATGAACCAAAAGTAAAAAATAAAAA 784
Db 747 TTTTATGAAATGAATGAACTGAAAAAACTTTGAAACGAAAAAATAAAAA 798

RESULT 6
AC012013 158198 bp DNA linear HTG 17-JUL-2001
LOCUS Homo sapiens chromosome 12 clone RP11-764L14, WORKING DRAFT
DEFINITION AC012013.18 GI:14717292
ACCESSION AC012013
VERSION AC012013
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Blinze,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burke,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douhwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frenzel,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garcia,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,C., Hart,M., Haylak,P., Hayes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C.,
Hollins,B., Homsli,F., Howard,S., Huber,J., Hulky,S., Hume,J.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvan,J.,
Kovar,C., Kratovic,U., Kuresh,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louiaged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc.help@bcm.tmc.edu
Project Information
Center project name: HMKZ
Center clone name: RP11-764L14
Summary Statistics
Sequencing vector: M13; 108821
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 161278 bases at least Q40
Consensus quality: 170806 bases at least Q30
Consensus quality: 174843 bases at least Q20
Estimated insert size: 168803; sum-of-coverage estimation
Estimated insert size: 163288; agarose-tp estimation
Quality coverage: 5.8x in Q20 bases; sum-of-coverage estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 108229: contig of 108229 bp in length
* 108230 108329: gap of unknown length
* 108330 150246: contig of 41917 bp in length
* 150247 150346: gap of unknown length
* 150347 153410: contig of 3064 bp in length
* 153411 153510: gap of unknown length
* 153511 153739: contig of 2229 bp in length
* 153740 155839: gap of unknown length
* 155840 158198: contig of 2359 bp in length.
Location/Qualifiers
1. 158198
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-764L14"
BASE COUNT 46057 a 30922 c 31747 g 49030 t 442 others
ORIGIN
Query Match 73.3%; Score 590.4; DB 2: Length 158198;
Best Local Similarity 87.0%; Pred. No. 3.2e-141;

```

[illegible]

FEATURES	source
<p><b>AUTHORS</b> Shepherd, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J., and O'Hara, P.</p> <p><b>TITLE</b> Direct Submission</p> <p><b>JOURNAL</b> Submitted (14-Sep-1999) Biomolecular Informatics, ZymoGenetics, Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA</p>	
<b>gene</b>	<p>1. .1017</p> <p>/organism="Mus musculus"</p> <p>/db_xref="dbEST:AA049839"</p> <p>/db_xref="taxon:10090"</p> <p>/clone="IMAGE:476061"</p> <p>1. .1017</p> <p>/gene="Zs19"</p> <p>343. . .891</p> <p>/gene="Zs19"</p> <p>/codon_start=1</p> <p>/product="putative secreted protein ZS19"</p> <p>/protein_id="AA01433.1"</p> <p>/db_xref="GI:6014636"</p> <p>/translation="MKGMGMLALLGLLTGTAHARSODLHCGARALVDELEIARVDPKRTIOMGSFRLNPDGOSVVEPVAPRAHAIETLEIFVDCDMKRYGQIDPSTHARKNYRVRVSRNGSESLDGLGIRIDSDLSGLTKFACISIVEIEDLEIFRSREADNVKDKLSKRTDLCDAHLHRSDEL"</p>
<b>CDS</b>	
<b>BASE COUNT</b>	248 a 250 c 315 g 204 t
<b>ORIGIN</b>	
Query Match	72.4%; Score 583.2; DB 10; Length 1017;
Best Local Similarity	88.3%; Pred. No. 8e-140;
Matches 657; Conservative 1; Mismatches 79; Indels 7; Gaps 2;	
QY 18	CAAAAGTAAAGTCCAGCGCGTCTGCCAGCGCTTGGGCCACGCGCGCGCCCTGGGACCAAA 77
Db 257	CAAAAGTGAAGTCCGGCGCTTCCAGAGCCCTGGGCCACGCGCGCGCGCTGGGACGAGA 316
QY 78	GGTGAGCAACCCCGTTCCTTAAARATGAAGAGCTGGGGTGGCTGGCGCTCGCTTCG 137
Db 317	GGTGAGCGAGCCCTGTACATAAAGATGAAGAGCTGGGGTGGCTGGCTTACTTTGG 376
QY 138	GGGCGCTCTGGGAAACGCGCTGGGCTCGGAGAGCCAGATCTCCACTGTGAGACATGCA 197
Db 377	GGGTCTCTCTGGGAACCTCCGCGCTCGAGAGAGCCAAAGATCTACACTGTGAGCTTGCA 436
QY 198	GGGCTCTGTGATGAATAGAAATGGGAAATTTGCCAGGTGGAGACCCCAAGAAACCATTC 257
Db 437	GGGCTCTGTGATGAATAGATGGGAAATTTGCCCGCTGGAGCCCAAGAAACCATTC 496
QY 258	AGATGGGATCTTTCGCGATCAATCCAGATGGCAGCGACCACTAGTGGTGGAGGTGCTTATG 317
Db 497	AGATGGGATCTTTCGCGATCAATCCAGATGGCAGCGACCACTAGTGGTGGAGGTGCTTATG 556
QY 318	CCCCGTCAGAGCCCCCACTCACAGAGCTGCTGGAGAGATATGTGACCGGATGAAGAGT 377
Db 557	CCCCGTCAGAGCCCCCACTCACAGAGGTGCTGGAGAGTGTGACCGAATGAAGAGT 616
QY 378	ATGGGGAACAGATTGATCTTCCACCCATCCCAAGAACTAGTACTAGTGGGCGGA 437
Db 617	ACGGGAACAGATTGATCTTCCACCCATCCCAAGAACTAGTACTAGTGGGCGGA 676
QY 438	ATGGGAATCCAGTGAACCTGACCTACAGGCATCCGATCGACTCAGATATTTAGCGCA 497
Db 677	ATGGGAATCCAGTGAACCTGACCTACAGGCATCCGATCGACTCAGATATTTAGCGCA 736
QY 498	CCCTCAAGTTTGGCGTGGAGAGCATTTGGAGGAATACGAGATGAAGAACTCATTTGATCT 557
Db 737	CCCTCAAGTTTGGCGTGGAGAGCATTTGGAGGAATACGAGATGAAGAACTCATTTGATCT 796
QY 558	TTTCCGAGAGGCTGACATGTTAAAGCAAACTTTCAGTACGAGCAAGATCTTTGTG 617
Db 797	TTTCCGAGAGGCTGACATGTTAAAGCAAACTTTCAGTACGAGCAAGATCTTTGTG 856
QY 618	ACCATGCCCCACATATCCGATGATGAGCTATGACACACTGGAGACGCCAC-----AC 672
Db 857	ACCATGCCCCACATATCCGATGATGAGCTATGACACACTGGAGACGCCAC-----AC 916

OY	673	TGGCTGTTATGCATACACCCCGCAGAGGGCAAAATGCTGTTCATTATATATATGCT	732
Db	917	AAAGTATATGACACACCCECCAGAGGGGAAGAATGACGACTTGCTTT--TATAATTACGT	974
OY	733	TTTTTACTGAAATTAACGTAAAAA	756
Db	975	TTTTTATGGAATGAACGTAAAAA	998
RESULT #			
AF186113		649 bp	mRNA
LOCUS			linear
DEFINITION	Homo sapiens putative secreted protein ZS1G9 (ZS1G9) mRNA, complete cds.		PRI 13-JAN-2000
ACCESSION	AF186113		
VERSION	AF186113.1	GI:6014631	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 649)		
AUTHORS	Sheppard,P., Jelinek,L., Whitmore,T., Blumberg,H., Lehner,J. and O'Hara,P.		
TITLE	Homo sapiens putative secreted protein		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 649)		
AUTHORS	Sheppard,P., Jelinek,L., Whitmore,T., Blumberg,H., Lehner,J. and O'Hara,P.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-SEP-1999) Biomolecular Informatics, ZymoGenetics, Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA		
FEATURES			
source	location/Qualifiers		
	1..649		
	/organism="Homo sapiens"		
	/db_xref="dbEST:R15306"		
	/db_xref="taxon:9606"		
	/chromosome="12"		
	/map="12015"		
	/clone="IMAGE:29675"		
gene	1..649		
	/gene="ZS1G9"		
CDS	104..358		
	/gene="ZS1G9"		
	/codon_start=1		
	/product="putative secreted protein ZS1G9"		
	/protein_id="AA01431.1"		
	/db_xref="GI:6014632"		
	/translation="MKGMGLALLALGLCTAMARRSODLHCACGRALVDELEWELIQ VDPKRTIOMSFRRNPDGSGSYEVTVTPNRYAHSGRG"		
BASE COUNT	175 a 150 c 171 g 149 t		4 others
ORIGIN			
Query Match	38.3%; Score 308.6;	DB 9;	Length 649;
Best Local Similarity	98.1%; Pred. No. 6,4e-69;		
Matches 312;	Conservative 0;	Mismatches 6;	Indels 0;
Gaps			0;
OY	1	CGGCCCAAGCGCTGGGGCCCAAAAGTAGAAATCCAGGCGTGTGCCACGCGCTTGGGCCACGCGC	60
Db	1	CGGCCCAAGCGCTGGGGCCCAAAAGTAGAAATCCAGGCGTGTGCCACGCGCTTGGGCCACGCGC	60
OY	61	CGCGCCCTGGGACCAAAGTAGAGACAACCCCGTTAACCTTAARATGAAGAAGCTGGGGTTG	120
Db	61	CGCGCCCTGGGACCAAAGTAGAGACAACCCCGTTAACCTTAARATGAAGAAGCTGGGGTTG	120
OY	121	GCTGGCCCTGCTTGTGGGGGCCCTGCTGCGGAACGCGCTGGGCTCGAGAGCCAGAGATCT	180
Db	121	GCTGGCCCTGCTTGTGGGGGCCCTGCTGCGGAACGCGCTGGGCTCGAGAGCCAGAGATCT	180
OY	181	CCACGTGGAGACATGACAGGCGCTGTGTGATGAGACTAGACTGAATGAGAAAATTGCCACAGTGA	240
Db	181	CCACGTGGAGACATGACAGGCGCTGTGTGATGAGACTAGACTGAATGAGAAAATTGCCACAGTGA	240

QY	241	CCCCAAGAGACCAATTCAGATGGGATCTTTCCGGATCAATCCAGATGGCAGCCACTACGT	300
Db	241	CCCCAAGAGACCAATTCAGATGGGATCTTTCCGGATCAATCCAGATGGCAGCCACTACGT	300
QY	301	GGTGAGAGTGCCTTATGC	318
Db	301	GGTGAGAGTGCCTTATGC	318
RESULT 9			
LOCUS	BC001027	824 bp	mRNA linear PRI 12-JUL-2001
DEFINITION	Homo sapiens, transmembrane protein 4, clone MGC:1545		
ACCESSION	BC001027	IMAGE:3344788, mRNA, complete cds.	
VERSION	BC001027.1	GI:12654402	
KEYWORDS	MGC.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 824)		
TITLE	Strausberg, R.		
JOURNAL	Direct Submission		
REMARK	Submitted (17-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
COMMENT	<p>NIH-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a></p> <p>Contact: MGC help desk</p> <p>Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a></p> <p>Tissue Procurement: ATCC</p> <p>cDNA Library Preparation: Rubin Laboratory</p> <p>cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)</p> <p>DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),</p> <p>Gaithersburg, Maryland:</p> <p>Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a></p> <p>Contact: <a href="mailto:nisc_mgc@nsl.nih.gov">nisc_mgc@nsl.nih.gov</a></p> <p>Shevchenko, Y., Weltherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantirip, S., Thomas, P.J., Tlionson, E.E., Touchman, J.W., Tsurgeson, C., Vogt, D.L., Walker, M.A., Zhang, L., H. and Green, E.D.</p> <p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a></p> <p>Series: IRAL Plate: 5 Row: P Column: 13</p> <p>This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6014631.</p> <p>location/Qualifiers</p> <p>1. 824</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="locustid:10330"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="MGC:1545 IMAGE:3344788"</p> <p>/tissue_type="Eye, retinoblastoma"</p> <p>/clone_id="NIH_MGC_16"</p> <p>/lab_host="DH10B-R"</p> <p>/note="Vector: pOTB7"</p> <p>250. 504</p> <p>/codon_start=1</p> <p>/product="transmembrane protein 4"</p> <p>/protein_id="AAH01027.1"</p> <p>/db_xref="GI:12654403"</p> <p>/translation="MKGSMALLILGTAARBSODLHCAGCRALVDELEMEINQVDPKRTQMSFRINPGSSVSFVYTVTPPNKVAHSGFG"</p>		

Query Match 35.8%; Score 288.4; DB 9; Length 824;  
 Best Local Similarity 95.8%; Pred. No. 1.1e-63;  
 Matches 295; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

11 CTTGGGCCCCAAATGAAAGTCCAGCGCTGCGAGCGCTGGGCGCCAGCGGGGGCCCTGG 70  
 |||||||  
 157 CTTGGGCCCCAAATGAAAGTCCAGCGCTGCTCCAGCGCTGGGCGCCAGCGGGGGCCCTGG 216  
 |||||||  
 71 GACCAAGGTGGAGCAACCCCGCTTACCCCTTAARATGAAAGGCTGGGGTGGCTGGCCCTG 130  
 |||||||  
 217 GAGCAGAGGTGGAGCGAGCCCATTAACGTAAGATGAAGAGCGTGGGCTGGGCGCCCTG 276  
 |||||||  
 131 CTTTGGGGGGCCCTGCTGGGAACCCGCTGGCTCGAGAGACCCAGAGATCTCCACTGTGGA 190  
 |||||||  
 277 CTTTGGGGGGCCCTGCTGGGAACCCGCTGGCTCGAGAGAGCCAGAGATCTCCACTGTGGA 336  
 |||||||  
 191 GCATCAGAGGCTCTGGTGGATGACAGAAATGGAAATGGCCAGGTGAGCCCAAGAG 250  
 |||||||  
 337 GCATCAGAGGCTCTGGTGGATGACAGAAATGGAAATGGCCAGGTGAGCCCAAGAG 396  
 |||||||  
 251 ACCATTACAGATGGATCTTCCGATCAATCCAGATGGCAGCCAGTCACTGTGAGAGTG 310  
 |||||||  
 397 ACCATTACAGATGGATCTTCCGATCAATCCAGATGGCAGCCAGTCACTGTGAGAGTG 456  
 |||||||  
 311 CCTTATGC 318  
 |||  
 457 ACTGTAC 464  
 |||

## RESULT 10

AX440375 564 bp DNA linear PAT 28-JUN-2002  
 LOCUS AX440375  
 DEFINITION Sequence 228 from Patent WO0190154.  
 ACCESSION AX440375  
 VERSION AX440375.1 GI:21651185  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Xu, J., Mitcham, J. L., Harlocker, S. L., Dillon, D. C., Secrist, H.,  
 Lodes, M. J., Algate, P. A., Pfling, S. P., Mannion, J., Benson, D. R. and  
 Carter, D.

TITLE Compositions and methods for the therapy and diagnosis of ovarian  
 cancer  
 JOURNAL Patent: WO 0190154-A 228 29-NOV-2001;  
 CORIXA CORPORATION (US)

FEATURES  
 source  
 1. 564  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

BASE COUNT 147 a 142 c 152 g 118 t 5 others  
 ORIGIN

Query Match 29.5%; Score 237.8; DB 6; Length 564;  
 Best Local Similarity 92.4%; Pred. No. 1.2e-50;

Matches 281; Conservative 1; Mismatches 13; Indels 9; Gaps 3;

11 CTTGGGCCCCAAATGAAAGTCCAGCGCTGCGAGCGCTGGGCGCCAGCGGGGGCCCTGG 70  
 |||||||  
 269 CTTGGGCCCCAAATGAAAGTCCAGCGCTGCTCCAGCGCTGGGCGCCAGCGGGGGCCCTGG 328  
 |||||||  
 71 GACCAAGGTGGAGCAACCCCGCTTACCCCTTAARATGAAAGGCTGGGGTGGCTGGCCCTG 130  
 |||||||  
 329 GAGCAGAGGTGGAGCGAGCCCATTAACGTAAGATGAAGAGCGTGGGCTGGGCGCCCTG 388  
 |||||||  
 131 CTTTGGGGGGCCCTGCTGGGAACCCGCTGGCTCGAGAGACCCAGAGATCTCCACTGTGGA 190  
 |||||||  
 389 CTTTGGGGGGCCCTGCTGGGAACCCGCTGGCTCGAGAGAGCCAGAGATCTCCACTGTGGA 447  
 |||||||  
 191 GCATCAGAGGCTCTGGTGGATGACAGAAATGGAAATGGCCAGGTGAGCCCAAGAG 250  
 |||||||

DB 448 -----AGGCGCTGCTGGTGGATGACAGAAATGGAAATGG-CCANGTGAGCCCAAGAG 500  
 QY 251 ACCATTACAGATGGATCTTCCGATCAATCCAGATGGCAGCCAGTCACTGTGAGAGTG 309  
 |||||||  
 DB 501 ACCATTACAGATGGATCTTCCGATCAATCCAGATGGCAGCCAGTCACTGTGAGAGTG 560  
 |||||||  
 QY 310 GCCT 313  
 |||  
 DB 561 GCCT 564  
 |||

## RESULT 11

AC025574 155023 bp DNA linear HTG 09-MAY-2002  
 LOCUS AC025574/c  
 DEFINITION Homo sapiens chromosome 12 clone RP11-348M3, WORKING DRAFT  
 AC025574  
 AC025574 12 unordered pieces.  
 AC025574.13 GI:20428723  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
 KEYWORDS  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 155023)  
 Muzny, D. M., Adams, C., Adio-Oduola, B., Ali-Osman, F. R., Allen, C.,  
 Alsbrooks, S. L., Amaral, H. C., Are, J. R., Ayala, M., Banks, T.,  
 Barbara, J., Benton, J., Blum, K., Blankenburg, K., Bonin, D.,  
 Bouck, J., Boyle, S., Brileva, M., Brown, M., Brown, N., Bryant, N. P.,  
 Buhay, C., Burch, P., Burkett, C., Burrell, K. L., Byrd, N. C.,  
 Carron, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,  
 Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R.,  
 Davila, M. L., Davis, C., Davy-Carroll, L., Dederich, D. A.,  
 Delaney, K. R., Delgado, O., Denn, A. L., Ding, Y., Dinh, H. H.,  
 Douthwaite, K. J., Draper, H., Dugan-Rocha, S., Durbin, K. J.,  
 Earnhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hogue, M., Hollway, C., Hollins, B.,  
 Homsl, F., Howard, S., Huber, J., Hulys, S., Hume, J., Jackson, L. E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jollivet, S., Joudah, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvan, J., Kovar, C.,  
 Kratovic, J., Kulesh, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L.,  
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louisse, H.,  
 Lohado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
 Massey, E., Mawhinney, E., McLeod, M. P., Meador, M., Mei, G., Metker, M.,  
 Miner, G., Miner, Z., Mitchell, J., Newton, N., Nguyen, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokkenko, S., Ogih, M., Okunolu, G.,  
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pu, L. L., Qules, M., Ren, Y.,  
 Rivers, M., Rojas, A., Rojudo, K., Ruff, M., Ruiz, S., Severy, G.,  
 Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,  
 Sodergren, E., Sonalike, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Tabor, P., Tameis, A., Tameis, K., Tang, H.,  
 Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wellington, S.,  
 Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G. and Glbbs, R.

Direct Submission

Unpublished

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

Submitted (11-MAR-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 155023)

AUTHORS  
TITLE  
JOURNAL

Worley, K.C.  
Direct Submission  
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On May 3, 2002 this sequence version replaced g1:20335511.

## COMMENT

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
Project information  
Center project name: HALY  
Center clone name: RP11-348M3  
Summary Statistics  
Sequencing vector: M13  
Chemistry: Dye-terminator Big Dye 950 of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 141522 bases at least Q40  
Consensus quality: 147097 bases at least Q30  
Consensus quality: 150442 bases at least Q20  
Estimated insert size: 150706; sum-of-coverage estimation  
Quality coverage: 4x in Q20 bases; sum-of-coverage estimation

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of 'N', but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 3033: contig of 3033 bp in length  
3034 3133: gap of unknown length  
3134 5971: contig of 2838 bp in length  
5972 6071: gap of unknown length  
6072 10269: contig of 4198 bp in length  
10270 10369: gap of unknown length  
10370 13009: contig of 2640 bp in length  
13010 13109: gap of unknown length  
13110 17573: contig of 4464 bp in length  
17574 17673: gap of unknown length  
17674 24650: contig of 6977 bp in length  
24651 24750: gap of unknown length  
24751 31778: contig of 7028 bp in length  
31779 31878: gap of unknown length  
31879 42263: contig of 10385 bp in length  
42264 42363: gap of unknown length  
42364 51104: contig of 8741 bp in length  
51105 51204: gap of unknown length  
51205 68306: contig of 17102 bp in length  
68307 68406: gap of unknown length  
68407 109481: contig of 41075 bp in length  
109482 109581: gap of unknown length  
109582 155023: contig of 45442 bp in length.

## FEATURES

Location/Qualifiers

1..155023  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/clone="RP11-348M3"

BASE COUNT 40008 a 36207 c 35987 g 41693 t 1128 others

## ORIGIN

Query Match 25.6%; Score 206; DB 2; Length 155023;  
Best Local Similarity 100.0%; Pred. No. 6.4e-42;  
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

306 AGGTGCTTATGCCCGCTCAGAGCCCACTCAGAGCTGCTGGAGGAATATGTGACC 365  
|||||

Db 56122 AGGTGCTTATGCCCGCTCAGAGCCCACTCAGAGCTGCTGGAGGAATATGTGACC 56063  
Oy 366 GGATGAAGAGATATGGGACAGATTTATCTTCACCACTCGCAAGACTAGTACTG 425  
Db 56062 GGATGAAGAGATATGGGACAGATTTATCTTCACCACTCGCAAGACTAGTACTG 56003  
Oy 426 TAGTGGCCGCGAATGAGAAATCCAGTACTGACCTACAGAGCATCGAATGACTAG 485  
Db 56002 TAGTGGCCGCGAATGAGAAATCCAGTACTGACCTACAGAGCATCGAATGACTAG 55943  
Oy 486 ATATTACCGCACCCTCAATTGGC 511  
Db 55942 ATATTACCGCACCCTCAATTGGC 55917

## RESULT 12

AC073896/c

LOCUS

DEFINITION

SEQUENCE, 5 unordered pieces.

AC073896

AC073896.37 GI:22450359

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

1 (bases 1 to 182350)

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.



COMMENT  
Baylor Plaza, Houston, TX 77030, USA  
On Aug 4, 2002 this sequence version replaced g1:22094227.

Center: Baylor College of Medicine  
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>  
Drafting Center Code: BCM

Contact: hgsc-help@bcm.tmc.edu  
Project Information

Center project name: HAFc  
Center clone name: RP11-183H16

Sequencing vector: Plasmid;  
Sequencing vector: M13;

Chemistry: Dye-terminator Big Dye 3; 98% of reads  
Assembly program: Phrap; version 0.990329

Consensus quality: 183845 bases at least Q40  
Consensus quality: 184849 bases at least Q30

Consensus quality: 185447 bases at least Q20  
Estimated insert size: 159026; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 9 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 2040: contig of 2040 bp in length  
\* 2041 2140: gap of unknown length  
\* 2141 4170: contig of 2030 bp in length  
\* 4171 4270: gap of unknown length  
\* 4271 8294: contig of 4024 bp in length  
\* 8295 8394: gap of unknown length  
\* 8395 14196: contig of 5802 bp in length  
\* 14197 14296: gap of unknown length  
\* 14297 33076: contig of 18780 bp in length  
\* 33077 33176: gap of unknown length  
\* 33177 52295: contig of 19119 bp in length  
\* 52296 52395: gap of unknown length  
\* 52396 84185: contig of 31790 bp in length  
\* 84186 84285: gap of unknown length  
\* 84286 120023: contig of 35738 bp in length  
\* 120024 120123: gap of unknown length  
\* 120124 184762: contig of 64639 bp in length.

FEATURES  
source  
1.184762  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/clone="RP11-183H16"

BASE COUNT 47425 a 44751 c 45101 g 46673 t 812 others

Query Match 25.6%; Score 206; DB 2; Length 184762;  
Best Local Similarity 100.0%; Pred. No. 6.7e-42;  
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 306 AGGTGCTTATGCGCCGCTCAGAGGCCACCTCAGAGCTGTGGAGATATGTGACC 365  
|||||  
Db 43484 AGGTGCTTATGCGCCGCTCAGAGGCCACCTCAGAGCTGTGGAGATATGTGACC 43543  
OY 366 GGATGAAGAGATAGGGAACAGATTGATCTTCCACCATCCGAGAACTACGTAGCTG 425  
|||||  
Db 43544 GGATGAAGAGATAGGGAACAGATTGATCTTCCACCATCCGAGAACTACGTAGCTG 43603  
OY 426 TAGTGGGCGGATGAGATCCAGTACCTGACCTACAGAGCATCCGATCGACTCAG 485  
|||||

Db 43604 TAGTGGGCGGATGAGATCCAGTACCTGACCTACAGAGCATCCGATCGACTCAG 43663  
OY 486 ATATTACGGCACCCTCAGATTGCG 511  
|||||  
Db 43664 ATATTACGGCACCCTCAGATTGCG 43689

RESULT 14  
AC090489

LOCUS 207424 bp DNA 11near ROD 26-JAN-2002  
DEFINITION Genomic sequence for Mus musculus, clone RP23-104010, complete  
sequence.

ACCESSION AC090489  
VERSION AC090489.8 GI:18376849  
KEYWORDS HTG.

SOURCE Mus musculus.  
ORGANISM Mus musculus.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;  
1 (bases 1 to 207424)

REFERENCE McCombie,W.R., de la Bastide,M., Spiegel,L., Preston,R.,  
Kirschhoff,K., Kuit,K., Nascimento,L., Zuber,T., Balla,V.,  
Bell,M., Baker,J., Santos,L., Miller,B., Katzenberger,F.,  
Muller,S., King,L., Yang,C., Palmer,L., O'Shaughnessy,A. and  
Dedhia,N.

TITLE Genomic sequence for Mus musculus, clone RP23-104010, complete  
sequence

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 207424)

AUTHORS McCombie,W.R.  
TITLE Direct Submission

JOURNAL Submitted (24-FEB-2001) Lita Annenberg Hazen Genome Sequencing  
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
Harbor, NY 11724, USA

REFERENCE 3 (bases 1 to 207424)  
AUTHORS McCombie,W.R.

TITLE Direct Submission  
JOURNAL Submitted (26-JAN-2002) Lita Annenberg Hazen Genome Sequencing  
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
Harbor, NY 11724, USA

REFERENCE 1.207424  
AUTHORS McCombie,W.R.

TITLE Direct Submission  
JOURNAL Submitted (26-JAN-2002) Lita Annenberg Hazen Genome Sequencing  
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
Harbor, NY 11724, USA

COMMENT On Jan 26, 2002 this sequence version replaced g1:18201765.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest.

FEATURES  
source  
1.207424  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-104010"  
/clone\_lib="RCI-23"  
/note="We believe the assembly to be correct. The sequence  
is a mononucleotide (T) repeat in which the exact number  
of Ts is unknown. Other subclones in the region show one  
fewer T than that represented in the assembly."

misc\_feature

BASE COUNT 52613 a 48671 c 49663 g 56477 t

Query Match 22.0%; Score 177.2; DB 10; Length 207424;  
Best Local Similarity 91.3%; Pred. No. 1.9e-34;  
Matches 188; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OY 306 AGGTGCTTATGCGCCGCTCAGAGGCCACCTCAGAGCTGTGGAGATATGTGACC 365  
|||||  
Db 75787 AGGTGCTTATGCGCCGCTCAGAGGCCACCTCAGAGCTGTGGAGATATGTGACC 75846  
OY 366 GGATGAAGAGATAGGGAACAGATTGATCTTCCACCATCCGAGAACTACGTAGCTG 425  
|||||  
Db 75847 GGATGAAGAGATAGGGAACAGATTGATCTTCCACCATCCGAGAACTACGTAGCTG 75906

QY 426 TACTGGCCGCGATGAGATCCAGTGAACCTGAGCAGTCCAGTCCAGTCCAGTCCAG 485  
 DB 75907 TCGTGGCCGCGATGAGATCCAGTGAACCTGAGCAGTCCAGTCCAGTCCAGTCCAG 75966  
 QY 486 ATATTAGCGCGACCCCTCAGTTTGGC 511  
 DB 75967 ATATCAGCGCGACCCCTCAGTTTGGC 75992

RESULT 15  
 AC109891 101666 bp DNA linear HNG 13-JUL-2002  
 LOCUS Rattus norvegicus clone CH230-31986, \*\*\* SEQUENCING IN PROGRESS  
 DEFINITION \*\*\* 47 unordered pieces.  
 AC109891  
 VERSION AC109891.3 GI:21738196  
 KEYWORDS HNG; HNGS PHASE1.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Ratus.

REFERENCE 1 (bases 1 to 101666)

AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayala,M., Banks,T.,  
 Barbada,J., Benton,J., Blomg,K., Blankenburg,K., Bonin,D.,  
 Bouck,J., Bowe,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,C., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earhart,E., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
 Homi,F., Howard,S., Huber,J., Huylk,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kratochvic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissege,D.,  
 Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzer,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokwaku,S., Ogih,M., Okunou,G.,  
 Ogunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Qules,M., Ren,Y.,  
 Rives,M., Rojas,A., Rojubenok,I., Rolfe,M., Ruiz,S., Savery,G.,  
 Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I.,  
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Syatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,O.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williamson,A., Wleczek,R., Woodem,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G. and Gldbs,R.  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 3 (bases 1 to 101666)

AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Worley,K.C.  
 Direct Submission  
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 12, 2002 this sequence version replaced g1.18860222.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GSKY  
 Center clone name: CH230-31986  
 ----- Summary Statistics  
 Sequencing vector: Plasmid  
 Chemistry: Dye-terminator Big Dye 100% of reads  
 Assembly program: Phrap version 0.990329  
 Consensus quality: 60229 bases at least Q40  
 Consensus quality: 63959 bases at least Q30  
 Consensus quality: 66730 bases at least Q20  
 -----

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
 \* NOTE: This is a working draft sequence. It currently  
 \* consists of 47 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1  
 \* 1423: contig of 1423 bp in length  
 \* 1424 1523: gap of unknown length  
 \* 1524 2972: contig of 1449 bp in length  
 \* 2973 3072: gap of unknown length  
 \* 3073 4717: contig of 1645 bp in length  
 \* 4718 4817: gap of unknown length  
 \* 4818 6632: contig of 1815 bp in length  
 \* 6633 6732: gap of unknown length  
 \* 6733 8040: contig of 1308 bp in length  
 \* 8041 8140: gap of unknown length  
 \* 8141 10043: contig of 1993 bp in length  
 \* 10044 10143: gap of unknown length  
 \* 10144 11337: contig of 1094 bp in length  
 \* 11338 11337: gap of unknown length  
 \* 11338 12362: contig of 1025 bp in length  
 \* 12363 12462: gap of unknown length  
 \* 12462 14100: contig of 1638 bp in length  
 \* 14101 14200: gap of unknown length  
 \* 14201 15840: contig of 1640 bp in length  
 \* 15841 15940: gap of unknown length  
 \* 15941 17053: contig of 1113 bp in length  
 \* 17054 17153: gap of unknown length  
 \* 17154 18836: gap of 1663 bp in length  
 \* 18837 18936: gap of unknown length  
 \* 18937 20137: contig of 1201 bp in length  
 \* 20138 20237: gap of unknown length  
 \* 20238 21736: contig of 1499 bp in length  
 \* 21737 21836: gap of unknown length  
 \* 21837 23519: contig of 1783 bp in length  
 \* 23520 23719: gap of unknown length  
 \* 23720 25907: contig of 2188 bp in length  
 \* 25908 26007: gap of unknown length  
 \* 26007 27406: contig of 1399 bp in length  
 \* 27407 27506: gap of unknown length  
 \* 27507 28837: contig of 1331 bp in length  
 \* 28838 28937: gap of unknown length  
 \* 28938 30329: contig of 1292 bp in length  
 \* 30330 30329: gap of unknown length  
 \* 30330 32012: contig of 1683 bp in length  
 \* 32013 32112: gap of unknown length  
 \* 32113 33198: contig of 1086 bp in length



```

* 33199 33298: gap of unknown length
* 33299 35519: contig of 2221 bp in length
* 35520 35619: gap of unknown length
* 35620 37677: contig of 2058 bp in length
* 37678 37777: gap of unknown length
* 37778 38936: contig of 1159 bp in length
* 38937 39036: gap of unknown length
* 39037 40273: contig of 1237 bp in length
* 40274 40373: gap of unknown length
* 40374 41906: contig of 1533 bp in length
* 41907 42006: gap of unknown length
* 42007 44506: contig of 2500 bp in length
* 44507 44606: gap of unknown length
* 44607 46803: contig of 2197 bp in length
* 46804 46903: gap of unknown length
* 46904 49226: contig of 2323 bp in length
* 49227 49326: gap of unknown length
* 49327 51433: contig of 2107 bp in length
* 51434 51533: gap of unknown length
* 51534 53763: contig of 2230 bp in length
* 53764 53863: gap of unknown length
* 53864 56613: contig of 2750 bp in length
* 56614 56713: gap of unknown length
* 56714 58773: contig of 2060 bp in length
* 58774 58873: gap of unknown length
* 58874 61527: contig of 2654 bp in length
* 61528 61627: gap of unknown length
* 61628 63192: contig of 1565 bp in length
* 63193 63292: gap of unknown length
* 63293 65362: contig of 2070 bp in length
* 65363 65462: gap of unknown length
* 65463 67438: contig of 1976 bp in length
* 67439 67538: gap of unknown length
* 67539 70505: contig of 2967 bp in length
* 70506 70605: gap of unknown length
* 70606 74524: contig of 3919 bp in length
* 74525 74624: gap of unknown length
* 74625 76737: contig of 2113 bp in length
* 76738 76837: gap of unknown length
* 76838 80133: contig of 3296 bp in length
* 80134 80233: gap of unknown length
* 80234 83122: contig of 2889 bp in length
* 83123 83222: gap of unknown length
* 83223 86731: contig of 3509 bp in length
* 86732 86831: gap of unknown length
* 86832 89929: contig of 3098 bp in length
* 89930 90029: gap of unknown length
* 90030 93107: contig of 3078 bp in length
* 93108 93207: gap of unknown length
* 93208 97419: contig of 4212 bp in length
* 97420 97519: gap of unknown length
* 97520 101666: contig of 4147 bp in length.

```

FEATURES  
source  
1.101666  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-319E6"

BASE COUNT 24650 a 22934 c 23569 g 23059 t 7454 others  
ORIGIN

Query Match 21.68; Score 174; DB 2; Length 101666;  
Best Local Similarity 90.38; Pred. No. 1.1e-33;  
Matches 186; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```

OY 306 AGGTGCTTATGCCGCTCAGAGCCGACCTCAGACAGAGCTGAGAGATATGTGACC 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 97570 AGGTTCTTATGCCGCTCAGAGCCGACCTCAGAGAGTGTGATGTGACC 97629
OY 366 GGATGAAGAGATGAGGAAACAGATTGCTTCACCCATCGCAGAACTACGACGTG 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 97630 GAATGAAGAGATGAGGAAACAGATTGCTTCACCCATCGCAGAACTACGACGTG 97689
OY 426 TAGTGGCCGGAAATGAGAACTCAGTGAACCTACAGGCAATCCGAATCGACTCAG 485

```

```

DB 97690 TTGTGCGCCGGAACGGAGATCCAGTGAAGCTAGACTTACAGGGTATCCGATTTCAG 97749
OY 486 ATATTAGCGCACCCCTCAAGTTTCCG 511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 97750 ATATCACTGGCACCCCTCAAGTTTCCG 97775

```

Search completed: March 30, 2003, 20:19:59  
Job time : 2658.55 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 17:35:17 ; Search time 191.291 Seconds  
(without alignments)  
9488.757 Million cell updates/sec

Title: US-10-082-502-16

Perfect score: 806  
Sequence: 1 cggcccaagctggggccaa.....aagagagagagagagacta 806

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database : N\_Geneseq.101002.\*  
1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	805.6	100.0	806	21	AAZ08293
2	804.6	99.8	806	20	AAZ06969
3	750.4	93.1	814	21	AAZ38327
4	750.4	93.1	814	24	ABK52765
5	750.4	93.1	814	24	ABL41995
6	750.4	93.1	814	24	ABK09772
7	749.4	93.0	1180	20	AAZ22112
8	749.4	93.0	1210	22	AAZ21330
9	702.4	87.1	714	20	AAZ97884

10	688.4	85.4	832	20	AAZ97837	Human secreted pro
11	646.4	80.2	657	24	ABK36007	CDNA sequence #398
12	599	74.3	1069	20	AAZ06970	Mouse secretory p3
13	599	74.3	1069	21	AAZ08294	Mouse ortholog gen
14	546	67.7	546	21	AAZ38326	Human transmembran
15	439.2	54.5	592	21	AAZ43525	Mouse secreted exp
16	334.8	41.5	558	22	AAZ18023	Human breast cancer
17	318.6	39.5	595	22	AAZ07882	Human breast cancer
18	308.6	38.3	649	20	AAZ06968	Human secretory pe
19	306.8	38.1	649	21	AAZ08284	Human Zs199 gene.
20	288	35.7	718	20	AAZ10655	CDNA encoding a hu
21	286	35.5	792	20	AAZ97836	Human secreted pro
22	284.2	35.3	1085	20	AAZ22130	WO9901020 Seq ID 3
23	279.2	34.6	649	20	AAZ97883	Human secreted pro
24	272.4	33.8	484	22	AAZ25628	Human breast cancer
25	271.4	33.7	1160	22	AAZ26689	Human breast cancer
26	267.6	33.2	415	20	AAZ06971	Zs199 expressed se
27	267	33.1	415	21	AAZ08285	Expressed Sequence
28	237.8	29.5	564	24	ABK09691	Human ovarian tumo
29	173.8	21.6	352	22	ABK09488	Human secreted pro
30	163.8	20.3	215	21	AAZ24962	Human secreted pro
31	151.8	18.8	265	20	AAZ39549	Human secreted pro
32	147	18.2	147	24	ABN94837	Gene #1335 used to
33	147	18.2	147	24	ABL65469	Lung cancer relate
34	130.8	16.2	495	22	AAZ180994	Human polynucleoti
35	126	15.6	126	24	ABN94836	Gene #1334 used to
36	126	15.6	126	24	ABL65468	Lung cancer relate
37	121	15.0	2532	22	ABA07341	Human pancreatic c
38	121	15.0	2532	22	AAZ32768	Human pancreatic c
39	121	15.0	2724	22	ABK07340	Human pancreatic c
40	121	15.0	2724	22	AAZ32767	Human genomic DNA
41	60	7.4	60	24	ABN39205	Human spliced tran
42	53.8	6.7	65	24	ABN55235	Mouse spliced tran
43	53.4	6.6	816	23	ABL21471	Drosophila melanog
44	51.4	6.4	3110	23	ABL21470	Drosophila melanog
45	50.4	6.3	483	23	ABV59155	Human prostate exp

## ALIGNMENTS

RESULT 1	AAZ08293	AAZ08293 standard; DNA; 806 BP.
ID	AAZ08293	
XX	AAZ08293	
AC	07-FEB-2000	(first entry)
XX		
DT	Human Zs199 gene encoding secretory protein variant-4.	
XX		
DE		
XX		
KW	Secretory protein-9; Human Zs199; chromosome 12q15 region; variant;	
KW	overexpression; antagonist; antibody; antisense nucleotide; tumour;	
KW	treatment; receptor; radio-label; fusion; polypeptide toxin; technique;	
KW	down-regulation; probe; diagnostic; therapeutic; cancer; brain; liver;	
KW	detection; stomach; lymphoma; alternative splicing; allelic variation;	
KW	silent mutation; ds.	
XX		
OS	Homo sapiens.	
XX		
XX	Synthetic.	
FT		
FT	Key	Location/Qualifiers
FT	CDS	104..652
FT		/tag= a
FT		/product= "Zs199 secretory protein variant-4"
FT		/note= "Overexpressed in tumours"
FT		/tag= b
FT	sig_peptide	104..163
FT		/tag= c
FT	mat_peptide	164..649
FT		/tag= c
FT		/label= Mature_Zs199_Protein_variant-4
PN	WO9960405-A1.	

XX 25-NOV-1999.  
 XX 19-MAY-1999; 99WO-US11107.  
 XX 19-MAY-1998; 98US-0081183.  
 XX (ZYMO ) ZYMOGENETICS INC.  
 XX Moore EE, Taft DW;  
 XX MPI: 2000-039447/03.  
 XX P-PSDB: AAI5135.  
 XX  
 PT Detecting tumors using antibodies, antagonists and antisense  
 PT nucleotides to secretory protein-9 (Zs19) -  
 PS  
 XX Disclosure; Page 33-35; 45pp; English.  
 CC The present DNA sequence is a gene encoding the variant of the secretory  
 CC protein-9, Zs19 that arises due to alternative splicing, allelic  
 CC variation or silent mutations that result in amino acid changes. This  
 CC sequence is mapped to the human chromosome 12q15 region. It is  
 CC overexpressed in tumors. Antagonists, antibodies and antisense  
 CC nucleotides to Zs19 are useful for detecting and treating tumors. The  
 CC antagonist may be an antibody or receptor to Zs19 and it may be radio-  
 CC labeled or fused to a polypeptide toxin. It can be used for down  
 CC regulating the overexpression of Zs19. The gene sequence can be used as  
 CC nucleic acid probes to detect RNA encoding Zs19. The Zs19 sequence  
 CC facilitates improved diagnostic and therapeutic techniques for detecting  
 CC and treating cancers, especially of the brain, liver, stomach, lymphoma,  
 CC etc., at an early stage.  
 XX  
 SQ Sequence 806 BP; 229 A; 183 C; 233 G; 160 T; 1 other;  
 Query Match 100.0%; Score 805.6; DB 21; Length 806;  
 Best Local Similarity 100.0%; Pred. No. 1,2e-193;  
 Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 481 CTCAGATATTAGCGGCACCCCTCAAGTTTGGCTGTGAGACATTGTGGAGAAATAGAGCA 540  
 OY 541 TGAACATCATTTGAATCTTTTCCGAGAGGCTGACAACTTTAAAGCAAACTTTGCAGTAA 600  
 DB 541 TGAACATCATTTGAATCTTTTCCGAGAGGCTGACAACTTTAAAGCAAACTTTGCAGTAA 600  
 OY 601 GCGAACAGATCTTTTGTGACCATGCCCTGCACATATTCGATGATGAGCTTTGAACACTGG 660  
 DB 601 GCGAACAGATCTTTTGTGACCATGCCCTGCACATATTCGATGATGAGCTTTGAACACTGG 660  
 OY 661 AGCAGCCACACTGGCTTATGATGATCACCCTCAGAGAGGAGAAATGTGCAATGCTTT 720  
 DB 661 AGCAGCCACACTGGCTTATGATGATCACCCTCAGAGAGGAGAAATGTGCAATGCTTT 720  
 OY 721 TATATATATATTTTACTGAAATTAACAGAAATATGAAACCAAAAGTAAAAA 780  
 DB 721 TATATATATATTTTACTGAAATTAACAGAAATATGAAACCAAAAGTAAAAA 780  
 OY 781 AAAAAAGAGAGAGAGAGAGAACTA 806  
 DB 781 AAAAAAGAGAGAGAGAGAACTA 806  
 RESULT 2  
 AAX06969  
 ID AAX06969 standard; cDNA; 806 BP.  
 AC AAX06969;  
 XX  
 AC 10-MAY-1999 (first entry)  
 DT  
 XX  
 DE Human secretory peptide-9 (Zs19) variant cDNA.  
 XX  
 KW Secretory peptide-9; Zs19; human; tumour marker; cancer; therapy;  
 KW diagnosis; growth enhancer; variant; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 104..652  
 FT /\*tag= a  
 FT sig\_peptide 104..163  
 FT /\*tag= b  
 FT mat\_peptide 164..649  
 FT /\*tag= c  
 XX  
 PN WO9901554-A1.  
 PD 14-JAN-1999.  
 XX  
 PF 02-JUL-1998; 98WO-US13859.  
 XX  
 PR 17-JUN-1998; 98US-0099005.  
 PR 03-JUL-1997; 97US-0051704.  
 PR 03-JUL-1997; 97US-0888088.  
 PR 19-MAY-1998; 98US-0081338.  
 PR 19-MAY-1998; 98US-0085983.  
 PR 17-JUN-1998; 98US-0089899.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Jaspers SR, Jelinek LJ, Sheppard PO, Whitmore TE;  
 XX  
 DR MPI: 1999-106055/09.  
 DR P-PSDB: AAM88474.  
 XX  
 PT New mammalian secretory peptide-9 (Zs19) - used as a growth  
 PT enhancer for placenta, liver and heart, and as an indicator of  
 PT cancer  
 XX  
 PS Claim 2; Page 73-74; 85pp; English.  
 XX  
 CC This cDNA clone encodes human secretory peptide-9, or Zs19,

CC variant (see AAW8469). Zs19 (see also AAW8469) is overexpressed in  
 CC human brain, liver, lung, oesophageal, stomach, colon, rectal,  
 CC thyroid and lymphoma tumors. Thus, Zs19 can be used as an  
 CC indicator for cancer. Zs19 cDNA was discovered in a placenta  
 CC clone from a full-term pregnancy cDNA library which contained an  
 CC expressed sequence tag (see AAW06971). The invention provides  
 CC polynucleotides (see AAW06968-70) encoding Zs19 polypeptides (see  
 CC AAW8469-77) including mature polypeptides, other processed forms,  
 CC variants and mouse orthologues. The Zs19 gene, or probes derived  
 CC from it, can be used to determine if Zs19 is present on chromosome  
 CC 12, and if a mutation has occurred. Antibodies raised against  
 CC Zs19 can be used as diagnostic agents to determine the presence of  
 CC Zs19, and thus the presence of cancer. They can also be labelled  
 CC with radioisotopes or fused with toxins and used to treat tumours  
 CC which overexpress Zs19. Antisense nucleotides derived from Zs19  
 CC cDNA can also be used to inhibit the growth of tumour cells. Zs19  
 CC proteins can be used to enhance the growth or development of the  
 CC placenta, heart or liver.

SQ Sequence 806 BP; 229 A; 183 C; 232 G; 160 T; 2 other;

Query Match 99.8%; Score 804.6; DB 20; Length 806;  
 Best Local Similarity 99.9%; Pred. No. 2.1e-193;

Matches 805; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGGCCCAAGGCTGGGGCCAAAGTCAAGCGGTCTGCCAGCGCTTGGGCGACGCG 60  
 DB 1 CGGCCCAAGGCTGGGGCCAAAGTCAAGCGGTCTGCCAGCGCTTGGGCGACGCG 60  
 OY 61 GGGGCGCTGGGACCAAGGTGGAGCAACCCGTTACCTTAATATAAAGGCTGGGCTTG 120  
 DB 61 GGGGCGCTGGGACCAAGGTGGAGCAACCCGTTACCTTAATATAAAGGCTGGGCTTG 120  
 OY 121 GGTGGCGCTGCTTCTGGGGGCGCTGCTGGGAACCGCGCTGGGCGCGAGAGCATCT 180  
 DB 121 GGTGGCGCTGCTTCTGGGGGCGCTGCTGGGAACCGCGCTGGGCGCGAGAGCATCT 180  
 OY 181 CCAGTGTGAGAGATGAGGAGGCTCTGTGTGATGAATAGATGGGAATTGCCAGGTGGA 240  
 DB 181 CCAGTGTGAGAGATGAGGAGGCTCTGTGTGATGAATAGATGGGAATTGCCAGGTGGA 240  
 OY 241 CCCCAAGAAAGACCATTCATGAGATGGGATCTTCCGGATCAATCCAGATGCCAGCATCA 300  
 DB 241 CCCCAAGAAAGACCATTCATGAGATGGGATCTTCCGGATCAATCCAGATGCCAGCATCA 300  
 OY 301 GGTGAGGTGCTTATGCCGCTCAGAGGCCCACTCACAGAGCTCTGGAGAGATATG 360  
 DB 301 GGTGAGGTGCTTATGCCGCTCAGAGGCCCACTCACAGAGCTCTGGAGAGATATG 360  
 OY 361 TGAACGGATGAAGAGATATGGGAGACAGATTGATCTTCCACCCATCGCAGAACTACGT 420  
 DB 361 TGAACGGATGAAGAGATATGGGAGACAGATTGATCTTCCACCCATCGCAGAACTACGT 420  
 OY 421 AGCTGTAGTGGGCGCGGAATGAGAAATCCAGTGAAGCTGACATCAAGGATCCGAATCGA 480  
 DB 421 AGCTGTAGTGGGCGCGGAATGAGAAATCCAGTGAAGCTGACATCAAGGATCCGAATCGA 480  
 OY 481 CTCAGATATTAGGCGCACCTCAAGTTTGCCTGTGAGAGCATTTGTGAGGAATACAGAGA 540  
 DB 481 CTCAGATATTAGGCGCACCTCAAGTTTGCCTGTGAGAGCATTTGTGAGGAATACAGAGA 540  
 OY 541 TGAATCATATGAATCTTTTCCGAGAGGCTGCAATGTTAAAGCAAACTTTGCAATGA 600  
 DB 541 TGAATCATATGAATCTTTTCCGAGAGGCTGCAATGTTAAAGCAAACTTTGCAATGA 600  
 OY 601 GCGAAGAGATCTTTGTGACATGCGCTGCACATATGCGATGATGACCTTGAACCACTGG 660  
 DB 601 GCGAAGAGATCTTTGTGACATGCGCTGCACATATGCGATGATGACCTTGAACCACTGG 660  
 OY 661 AGCAGCCCACTGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 DB 661 AGCAGCCCACTGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 720

OY 721 TATATATATATGTTTTTACTGAATTAAGTCAAAAAATATGAAACCAAAAGTAAAAAAA 780  
 DB 721 TATATATATATGTTTTTACTGAATTAAGTCAAAAAATATGAAACCAAAAGTAAAAAAA 780  
 OY 781 AAAAAAAGAGAGAGAGAGAGAACTA 806  
 DB 781 AAAAAAAGAGAGAGAGAGAACTA 806

# RESULT 3

AAZ38327  
 ID AAZ38327 standard; cDNA; 814 BP.

AAZ38327;

09-FEB-2000 (first entry)

Human transmembrane protein cDNA clone HP10390.

HP10390; transmembrane domain; stomach cancer cell; antibody;

assay reagent; diagnostic marker; primer; probe; antisense; gene therapy;  
 agonist; antagonist; ligand; therapeutic; ds.

Homo sapiens.

Key Location/Qualifiers

CDS 145..693  
 /ftg="a  
 /product="human transmembrane protein HP10390"

W09955862-A2.

04-NOV-1999.

27-APR-1999; 99W0-JP02226.

28-APR-1998; 98JP-0119395.

(SAGA) SAGAMI CHEM RES CENT.

(PROT-) PROTEGENE INC.

Kato S, Kimura T;

WPI: 2000-023358/02.

P-PSDB; AAY52391.

Human proteins with transmembrane domains, involved in control of cell  
 proliferation and differentiation, useful for treating e.g. cancer or  
 inflammation

Claim 4; Page 106-107; 114pp; English.

This sequence represents the human cDNA clone HP10390

which encodes a 20 kD protein with one putative transmembrane

domain in the N-terminus. The cDNA was isolated from a human stomach

cancer cell line cDNA library. The protein has no homology with any

known protein. The protein may be used to raise specific antibodies, as

assay reagents, as diagnostic tissue markers, for the isolation of

cognate receptors, ligands and binding proteins, and as biologically

active agents. Nucleotides encoding the protein may be used as primers

and probes or antisense molecules, and in gene therapy. Cells transformed

with these nucleotides may be used to screen for agonists and antagonists

which are potentially useful therapeutically.

Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Query Match 93.1%; Score 750.4; DB 21; Length 814;  
 Best Local Similarity 99.0%; Pred. No. 1.1e-179;

Matches 754; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 11 CTGGGGCCAAAGTGAAGTCAAGCGGTCTGCCAGCGCTTGGGCGACGCGGCGCCCTGG 70  
 DB 52 CTGGGGCCAAAGTGAAGTCAAGCGGTCTGCCAGCGCTTGGGCGACGCGGCGGCGCCCTGG 111

QY	71	GACCAAGGTGGAGCAACCCCGTTACCTTAAARATGAAAGCTGGGGTTGGCTGGCCCTT	130
Db	112	GAGCAGAGGTGGAGCAACCCCATTAACCTTAAAGATGAAGCTGGGGTTGGCTGGCCCTT	171
QY	131	CTTCTGGGGGGCCCTGCTGGGAACCGCCTGGCTCGAGAGACAGATCTCCACTGTGA	190
Db	172	CTTCTGGGGGGCCCTGCTGGGAACCGCCTGGGGCTCGAGAGACAGATCTCCACTGTGGA	231
QY	191	GCATGACAGGCTCTGGTGGATGAACCTGAATGGGAATTTGCCAGGTGGAGCCCAAGAA	250
Db	232	GCATGACAGGCTCTGGTGGATGAACCTGAATGGGAATTTGCCAGGTGGAGCCCAAGAA	291
QY	251	ACCAATTGATGGGATCTTTCCGGATCAATTCAGATGGCAGCCAGTCAGTGTTGGAGGTG	310
Db	292	ACCAATTGATGGGATCTTTCCGGATCAATTCAGATGGCAGCCAGTCAGTGTTGGAGGTG	351
QY	311	CCCTATGCCCGCTGAGAGGCCACCTCACAGAGTCTGTGAGAGAGATATGTACCCGATG	370
Db	352	CCCTATGCCCGCTGAGAGGCCACCTCACAGAGTCTGTGAGAGAGATATGTACCCGATG	411
QY	371	AAGGAGTATGGGGGAACGATTGATCCTTCCACCACATGCAAGAACTACGTACCTGTAGTG	430
Db	412	AAGGAGTATGGGGGAACGATTGATCCTTCCACCACATGCAAGAACTACGTACCTGTAGTG	471
QY	431	GCCCGGAATGGAAGATTCATGAACTGGACCTTACAAAGCATTCGGAATCGACTCAGATATT	490
Db	472	GCCCGGAATGGAAGATTCATGAACTGGACCTTACAAAGCATTCGGAATCGACTCAGATATT	531
QY	491	AGCGGCAACCTCAAGTTTGGCTGGAAGCATTTGTGGAGAAATACGAGATGTAACATTT	550
Db	532	AGCGGCAACCTCAAGTTTGGCTGGAAGCATTTGTGGAGAAATACGAGATGTAACATTT	591
QY	551	GAATTCCTTTCCGAGAGAGCTGACATGTTTAAAGACAAACTTTGCAGTAAGCAGACAGAT	610
Db	592	GAATTCCTTTCCGAGAGAGCTGACATGTTTAAAGACAAACTTTGCAGTAAGCAGACAGAT	651
QY	611	CTTTGTGACCAATGGCCGACATTTGTGCATGATGAGCTATGAACCACTGGAGGAGCCGAC	670
Db	652	CTTTGTGACCAATGGCCCTGCACTATTCGCAATGATGAGCTATGAACCACTGGAGGAGCCGAC	711
QY	671	ACTGGCTGATGATGCACCCCGACAGAGAGGGAATAATGGTGGCAATGCCCTTTATATATTAT	730
Db	712	ACTGGCTGATGATGCACCCCGACAGAGAGGGAATAATGGTGGCAATGCCCTTTATATATTAT	771
QY	731	GTTTTACTGAATTAACGTAAAAAATATATAAACCAAAAGTA	772
Db	772	GTTTTACTGAATTAACGTAAAAAATATATAAACCAAAAGTA	813

RESULT	4
ABK52765	
ID	ABK52765 standard; cDNA; 814 BP.
XX	
AC	ABK52765;
XX	
DT	15-AUG-2002 (first entry)
XX	
DE	cDNA encoding transmembrane protein 4, a cancer-linked protein.
XX	
KW	Expressed sequence tag; EST; human; cancer; anti-neoplastic;
KM	cytotoxic T lymphocyte; chemotherapy; cytostatic; chromosome 12;
KW	transmembrane protein 4; gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
CD5	145..693
FT	/tag= "a
XX	/product= "Transmembrane protein 4"
PX	
PN	
XX	
MO200231198-A2.	

Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
11	CTGGGGCCCAAGTGAATGCCAGCGGCTTGCCAGCGCTTGGGCCACGGCGGCGCCCTGCG	93.1%;	Score 750.4;	DB 24;	Length 814;	
52	CTGGGGCCCAAGTGAATGCCAGCGGCTTGCCAGCGCTTGGGCCACGGCGGCGCCCTGCG	99.0%;	Pred. No. 1.1e-179;			
71	GACCAAGGTGAGAACACCCCTTACCTTAAATGAAAGAGGTGGGGTTGGCTGGGCCCTG	93.1%;	Score 750.4;	DB 24;	Length 814;	
112	GAGCAGAGGTGAGAGGACCCCTTACCTTAAATGAAAGAGGTGGGGTTGGCTGGGCCCTG	99.0%;	Pred. No. 1.1e-179;			
131	CTTCTGGGGGCGCTCTGCTGGGAACCGCTGGGCTCGGAGGAGCCAGATCTCCACTGTGGA	93.1%;	Score 750.4;	DB 24;	Length 814;	
172	CTTCTGGGGGCGCTCTGCTGGGAACCGCTGGGCTCGGAGGAGCCAGATCTCCACTGTGGA	99.0%;	Pred. No. 1.1e-179;			
191	GCATCAGAGGCTCTGCTGGATGACTAGAAATGGAAATTTGCCAGGTGAGACCCCAAGANG	93.1%;	Score 750.4;	DB 24;	Length 814;	
232	GCATCAGAGGCTCTGCTGGATGACTAGAAATGGAAATTTGCCAGGTGAGACCCCAAGANG	99.0%;	Pred. No. 1.1e-179;			
251	ACCATTCAGATGGAGTCTTTCGGATCAATCAGATGAGCAGCACTCACTGTGAGGTG	93.1%;	Score 750.4;	DB 24;	Length 814;	
292	ACCATTCAGATGGAGTCTTTCGGATCAATCAGATGAGCAGCACTCACTGTGAGGTG	99.0%;	Pred. No. 1.1e-179;			

QY 311 CCTTATGCCCGCTCAGAGGCCACCTCAGACAGCTGCTGGAGAGATATGTACCGGATG 370  
DB 352 CCTTATGCCCGCTCAGAGGCCACCTCAGACAGCTGCTGGAGAGATATGTACCGGATG 411  
QY 371 AAGAGATATGGGAGACAGATGATCTCTCCACCATCGCAAGAACTACGATGATGATG 430  
DB 412 AAGAGATATGGGAGACAGATGATCTCTCCACCATCGCAAGAACTACGATGATGATG 471  
QY 431 GGGCGAATGAGAAATCAGATGAACTGACCTCAAGAGCATCCGAATCGACATGATAT 490  
DB 472 GGGCGAATGAGAAATCAGATGAACTGACCTCAAGAGCATCCGAATCGACATGATAT 531  
QY 491 AGCGGACCCCTCAAGTTTGGCTGTGAGAGCATTTGGAGAAATACGAGATGAATCTATT 550  
DB 532 AGCGGACCCCTCAAGTTTGGCTGTGAGAGCATTTGGAGAAATACGAGATGAATCTATT 591  
QY 551 GAATCTTTTCCGAGAGGGGTGACATGTTAAAGACAAACTTTGCACTAGCAACAGAT 610  
DB 592 GAATCTTTTCCGAGAGGGGTGACATGTTAAAGACAAACTTTGCACTAGCAACAGAT 651  
QY 611 CTCTGTGACCATGCGCTGACATATGCAATGATGATGATGATGATGATGATGATGAT 670  
DB 652 CTCTGTGACCATGCGCTGACATATGCAATGATGATGATGATGATGATGATGATGAT 711  
QY 671 ACTGCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730  
DB 712 ACTGCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771  
QY 731 GTTTTACTGAATTAAGTGAATAAATATGAACCAAAAGTA 772  
DB 772 GTTTTACTGAATTAAGTGAATAAATATGAACCAAAAGTA 813

RESULT 5  
ABLA1995  
ID ABLA1995 standard; DNA; 814 BP.  
AC ABLA1995;  
XX  
XX  
XX 11-JUN-2002 (first entry)  
DE Nucleotide sequence of human polypeptide HP10390.  
XX  
XX Antibody; antigen; transmembrane domain protein; HP10390; gene; ss.  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT CDS 145..693  
FT /tag= a  
FT /product= "HP10390"  
XX  
XX PN WO200208416-A1.  
XX  
XX PD 31-JAN-2002.  
XX  
XX PE 24-JUL-2001; 2001WO-JP06371.  
XX  
XX PR 24-JUL-2000; 2000JP-0222743.  
XX  
XX PR 24-AUG-2000; 2000JP-0254407.  
XX  
XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
XX  
XX PI Kato S, Nagata N, Fujimura N, Kobayashi M, Ito K, Ishizuka Y;  
XX  
XX DR P-PSDB; ABB09717.  
XX  
XX WPI: 2002-195877/25.  
XX  
XX DR P-PSDB; ABB09717.  
XX  
XX Antibody preparation by inoculation of an animal with a vector  
XX expressing a fusion protein of an antigen on the C-terminal side of a  
XX transmembrane domain for use as drugs, diagnostic reagents and  
XX laboratory reagents

PS Example; Page 33-35; 45pp; Japanese.  
XX  
XX The specification describes a method of antibody preparation. The  
CC method comprises inoculating an animal with a vector expressing  
CC a fusion protein having an antigen protein fused to the C-terminal  
CC side (extracellular) of a transmembrane domain protein (the  
CC N-terminal side of which is intracellular), and then isolating and  
CC purifying the antibody from the animal. The antibodies can be used  
CC as drugs, vaccines, diagnostic reagents and laboratory reagents. The  
CC present sequence encodes a polypeptide, designated HP10390, which was  
CC used in the course of the invention.

XX  
XX Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;  
SQ

Query Match 93.1%; Score 750.4; DB 24; Length 814;  
Best local similarity 99.0%; Pred. No. 1.1e-179;  
Matches 754; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 11 CTGGGCGCAAAAGTGAAGTCCAGCGGTCTCCAGCGCTTGGCCAGCGCGGCGCTG 70  
DB 52 CTGGGCGCAAAAGTGAAGTCCAGCGGTCTCCAGCGCTTGGCCAGCGCGGCGCTG 111  
QY 71 GACCAAAAGTGGAGCAACCCGTTACCTTAATATGAAGAGCTGGGGTGGTGGCCCTG 130  
DB 112 GACCAAAAGTGGAGCAACCCGTTACCTTAATATGAAGAGCTGGGGTGGTGGCCCTG 171  
QY 131 CTCTGTGAGGCGCTGCTGGAGAACCGGCTGGGCGGAGGAGGATCTCCACTGTGA 190  
DB 172 CTCTGTGAGGCGCTGCTGGAGAACCGGCTGGGCGGAGGAGGATCTCCACTGTGA 231  
QY 191 GCATGACAGGCGCTGCTGGAGAACCGGCTGGGCGGAGGAGGATCTCCACTGTGA 250  
DB 232 GCATGACAGGCGCTGCTGGAGAACCGGCTGGGCGGAGGAGGATCTCCACTGTGA 291  
QY 251 ACCATTCAGATGGAGATCTTCCGATCAATCCAGATGCGACCGCATGATGATGATG 310  
DB 292 ACCATTCAGATGGAGATCTTCCGATCAATCCAGATGCGACCGCATGATGATGATG 351  
QY 311 CCTTATGCCCGCTCAGAGGCCACCTCAGACAGCTGCTGGAGAGATATGTACCGGATG 370  
DB 352 CCTTATGCCCGCTCAGAGGCCACCTCAGACAGCTGCTGGAGAGATATGTACCGGATG 411  
QY 371 AAGAGATATGGGAGACAGATGATCTCTCCACCATCGCAAGAACTACGATGATGATG 430  
DB 412 AAGAGATATGGGAGACAGATGATCTCTCCACCATCGCAAGAACTACGATGATGATG 471  
QY 431 GGGCGAATGAGAAATCAGATGAACTGACCTCAAGAGCATCCGAATCGACATGATAT 490  
DB 472 GGGCGAATGAGAAATCAGATGAACTGACCTCAAGAGCATCCGAATCGACATGATAT 531  
QY 491 AGCGGACCCCTCAAGTTTGGCTGTGAGAGCATTTGGAGAAATACGAGATGAATCTATT 550  
DB 532 AGCGGACCCCTCAAGTTTGGCTGTGAGAGCATTTGGAGAAATACGAGATGAATCTATT 591  
QY 551 GAATCTTTTCCGAGAGGGGTGACATGTTAAAGACAAACTTTGCACTAGCAACAGAT 610  
DB 592 GAATCTTTTCCGAGAGGGGTGACATGTTAAAGACAAACTTTGCACTAGCAACAGAT 651  
QY 611 CTCTGTGACCATGCGCTGACATATGCAATGATGATGATGATGATGATGATGATGAT 670  
DB 652 CTCTGTGACCATGCGCTGACATATGCAATGATGATGATGATGATGATGATGATGAT 711  
QY 671 ACTGCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730  
DB 712 ACTGCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771  
QY 731 GTTTTACTGAATTAAGTGAATAAATATGAACCAAAAGTA 772  
DB 772 GTTTTACTGAATTAAGTGAATAAATATGAACCAAAAGTA 813

RESULT 6  
ABK09772

ID ABR09772 standard; cDNA; 814 BP.  
XX  
AC ABR09772;  
XX  
DT 14-MAR-2002 (first entry)  
XX  
DE Human ovarian tumour protein encoding cDNA #305.  
XX  
KM Human; ovarian tumour protein; cancer; cytostatic; immunostimulant; ss;  
KM gene therapy; CD4+ T cell; CD8+ T cell; PCR primer.  
XX  
OS Homo sapiens.  
XX  
PN WO200190154-A2.  
XX  
PD 29-NOV-2001.  
XX  
PF 23-MAY-2001; 2001WO-US16895.  
XX  
PR 24-MAY-2000; 2000US-207107P.  
PR 13-JUN-2000; 2000US-211457P.  
PR 21-JUN-2000; 2000US-213673P.  
PR 03-AUG-2000; 2000US-223288P.  
PR 01-MAR-2001; 2001US-272790P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Xu J, Mitcham JL, Harlocker SL, Dillon DC, Secrist H, Lodes MO;  
PI Algaete PA, Fling SP, Mannion J, Benson DR, Carter D;  
XX  
DR WPI; 2002-097641/13.  
XX  
PT New isolated polynucleotide encoding polypeptide comprising portion of  
PT ovarian tumour protein, useful for detection, diagnosis and therapy of  
PT human ovarian cancer.  
XX  
PS  
XX  
PS Claim 1; Page 253-254; 285pp; English.  
XX  
CC The invention relates to an isolated polynucleotide encoding a  
CC polypeptide comprising a portion of an ovarian tumour protein. The  
CC sequences of the invention are useful for stimulating an immune response  
CC and for treating ovarian cancer in a patient. An antigen presenting cell  
CC that expresses the sequences is useful for treating ovarian cancer by  
CC incubating CD4+ and/or CD8+ T cells isolated from a patient. The T cells  
CC can then be proliferated and administered to the patient to inhibit the  
CC development of cancer. The DNA sequences are useful as probes or primers  
CC for nucleic acid hybridisation, to direct expression of a polypeptide in  
CC appropriate host cells. Detecting the presence of a cancer in a patient  
CC involves obtaining a biological sample from the patient, contacting the  
CC biological sample with an agent that binds to the protein, detecting the  
CC amount of protein that binds to the agent, comparing the amount of  
CC protein to a predetermined cut-off value and determining the presence of  
CC cancer. Sequences ABR09464-ABR09802 represent PCR primers and cDNA  
CC molecules encoding ovarian tumour proteins of the invention.  
XX  
XX  
SO Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;  
Query Match 93.1%; Score 750.4; DB 24; Length 814;  
Best Local Similarity 99.0%; Pred. No. 1.1e-179;  
Matches 754; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 191 GCATCAGGCGCTCTGTGATGATAGACTAGATGGAAATTTGCCAGTGGACCCCAAGAAG 250  
DB 232 GCATCAGGCGCTCTGTGATGATAGACTAGATGGAAATTTGCCAGTGGACCCCAAGAAG 291  
QY 251 ACCATTTCAGATGGATGATCTTCCGATCAATCCAGATGGAGCCAGTGCATGATGGAGGG 310  
DB 292 ACCATTTCAGATGGATGATCTTCCGATCAATCCAGATGGAGCCAGTGCATGATGGAGGG 351  
QY 311 CCTATGCCCCGCTCAGAGGCCACCTCAGACAGTCTGTGAGAGATATGTGACCGGATG 370  
DB 352 CCTATGCCCCGCTCAGAGGCCACCTCAGACAGTCTGTGAGAGATATGTGACCGGATG 411  
QY 371 AAGGAGTATGGGGAACAGATTGATCTTCCAGATCCAGATGGAGCCAGTGCATGATGGAGGG 430  
DB 412 AAGGAGTATGGGGAACAGATTGATCTTCCAGATCCAGATGGAGCCAGTGCATGATGGAGGG 471  
QY 431 GCGCGGAATGGAGATCCAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATG 490  
DB 472 GCGCGGAATGGAGATCCAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATG 531  
QY 491 AGCGGACCCCTCAGATGATGCTGTGAGAGCATTTGTGAGGAATAGAGAGATGACTCAT 550  
DB 532 AGCGGACCCCTCAGATGATGCTGTGAGAGCATTTGTGAGGAATAGAGAGATGACTCAT 591  
QY 551 GAATCTTTTCCCGAGAGGCTGACATGTTTAAAGCAAACTTTGCACTAAGCGAAACAGAT 610  
DB 592 GAATCTTTTCCCGAGAGGCTGACATGTTTAAAGCAAACTTTGCACTAAGCGAAACAGAT 651  
QY 611 CTTGTGACATGACCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 670  
DB 652 CTTGTGACATGACCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 711  
QY 671 ACTGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 730  
DB 712 ACTGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 771  
QY 731 GTTTACTGGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 792  
DB 772 GTTTACTGGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 813  
RESULT 7  
ID AAX22112  
XX  
XX AAX22112;  
XX  
DT 18-MAY-1999 (first entry)  
XX  
DE Human secreted protein gene 2 clone H2MB56.  
XX  
XX Human; secreted protein; gene therapy; protein therapy; tissue; cancer;  
KM tumour; neurodegenerative disorder; leukaemia; autoimmune disease; AIDS;  
KM developmental abnormality; foetal deficiency; Alzheimer's disease;  
KM cognitive disorder; schizophrenia; immunological disorder; mood disorder;  
KM immune deficiency disease; respiratory disorder; arthritis; skeletal;  
KM haematopoietic disorder; neural; osteoporosis; metabolic disorders;  
KM cardiovascular; endocrine; gastrointestinal; asthma; diagnosis; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO9901020-A2.  
XX  
PD 14-JAN-1999.  
XX  
XX  
PF 30-JUN-1998; 98WO-US13608.  
XX  
XX  
PR 12-SEP-1997; 97US-0058663.  
PR 01-JUL-1997; 97US-0051381.  
PR 01-JUL-1997; 97US-0051480.  
PR 12-SEP-1997; 97US-0058598.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.



XX Carter KC, Endress GA, Feng P, Rosen CA, Ruben SM:  
PI WPI: 1999-105683/09.  
DR P-PSDB: AAY01136, AAY01162, AAY01163.  
XX  
PT New isolated human genes and the secreted polypeptides they encode -  
PT useful for diagnosis and treatment of e.g. cancers, neurological  
PT disorders, immune diseases, immune deficiency diseases or blood  
XX  
PS Claim 4: Page 117; 179pp: English.  
XX  
CC The invention relates to nucleic acid sequences (AA022111 to AA022124)  
CC encoding human secreted proteins (AAY01135 to AAY01158). The secreted  
CC protein gene sequences are deposited with the ATCC under deposit number  
CC ATCC 209118. Host cells comprising recombinant vectors containing the  
CC nucleic acid sequences are used for the recombinant production of the  
CC secreted proteins. The polynucleotide and amino acid sequences are useful  
CC for are useful for preventing, treating or ameliorating medical  
CC conditions e.g. by protein or gene therapy. Pathological conditions can  
CC be also diagnosed by determining the amount of the new polypeptides in a  
CC sample or by determining the presence of mutations in the new  
CC polynucleotides. Specific uses are described for each of the  
CC polynucleotides, based on which tissues they are most highly expressed  
CC in, and include developing products for the diagnosis or treatment of  
CC cancer, tumours, developmental abnormalities and foetal deficiencies,  
CC autoimmune diseases, lymphomas, Alzheimer's and cognitive disorders,  
CC schizophrenia, immunological disorders, immune deficiency diseases  
CC (AIDS), mood disorders, respiratory disorders, arthritis, asthma,  
CC haematopoietic disorders, neural disorders, skeletal disorders,  
CC osteoporosis, metabolic disorders, cardiovascular disorders, endocrine  
CC disorders or gastrointestinal disorders. The polypeptides are also useful  
CC for identifying their binding partners. The present sequence represents a  
CC gene encoding a human secreted protein (see descriptor line for gene  
CC number and clone identification).  
XX  
XX Sequence 1180 BP; 258 A; 304 C; 363 G; 250 T; 5 other:  
SQ  
Query Match 93.0%; Score 749.4; DB 20; Length 1180;  
Best Local Similarity 98.9%; Pred. No. 2,2e-179;  
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

DB 840 GGCCGGAAATGAGAAATCCAGTGAACCTGACCTACAAAGCAATCCGAATCAGATTAAT 899  
OY 491 AGCGGACCCCTCAAGTTCCTGCTGAGAGCAATTTGAGAGAAATACGAGATCAATCAAT 550  
DB 900 AGCGGACCCCTCAAGTTCCTGCTGAGAGCAATTTGAGAGAAATACGAGATCAATCAAT 959  
OY 551 GAATTCCTTCCCGAGAGGCTGACAAATGTTAAAGACAACTTTGACAGTAAAGCAAGAT 610  
DB 960 GAATTCCTTCCCGAGAGGCTGACAAATGTTAAAGACAACTTTGACAGTAAAGCAAGAT 1019  
OY 611 CTTTGGACCAATCCCTCGACATATGCAATGATGACCTATGAAACCACTGGAGACCCAC 670  
DB 1020 CTTTGGACCAATCCCTCGACATATGCAATGATGACCTATGAAACCACTGGAGACCCAC 1079  
OY 671 ACTGCGTATGATGATCAACCCCGAGAGGGAATAATGTCGCAATGCTTTATATATTAAT 730  
DB 1080 ACTGCGTATGATGATCAACCCCGAGAGGGAATAATGTCGCAATGCTTTATATATTAAT 1139  
OY 731 GTTTTACTGAAATTAAGTAAATATGAAACCAAAAGT 771  
DB 1140 GTTTTACTGAAATTAAGTAAATATGAAACCAAAAGT 1180  
RESULT 8  
ID AAS21330 standard; cDNA; 1210 BP.  
AC AAS21330;  
AC AAS21330;  
DT 24-OCT-2001 (first entry)  
XX  
DE Human cDNA sequence encoding for PRO4426 polypeptide.  
XX  
KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
KW breast; prostate; cervical; tumour necrosis factor- $\alpha$ ; TNF- $\alpha$ ;  
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200140466-A2.  
XX  
PD 07-JUN-2001.  
XX  
PF 01-DEC-2000; 2000MO-US32678.  
XX  
PR 01-DEC-1999; 99MO-US28301.  
PR 01-DEC-1999; 99MO-US28634.  
PR 02-DEC-1999; 99MO-US28551.  
PR 02-DEC-1999; 99MO-US28564.  
PR 02-DEC-1999; 99MO-US28565.  
PR 09-DEC-1999; 99MO-US28565.  
PR 16-DEC-1999; 99MO-US30911.  
PR 20-DEC-1999; 99MO-US30911.  
PR 20-DEC-1999; 99MO-US30911.  
PR 30-DEC-1999; 99MO-US30911.  
PR 30-DEC-1999; 99MO-US31243.  
PR 06-JAN-2000; 2000MO-US00277.  
PR 06-JAN-2000; 2000MO-US00376.  
PR 18-FEB-2000; 2000MO-US03565.  
PR 18-FEB-2000; 2000MO-US04341.  
PR 18-FEB-2000; 2000MO-US04342.  
PR 22-FEB-2000; 2000MO-US04414.  
PR 24-FEB-2000; 2000MO-US04914.  
PR 24-FEB-2000; 2000MO-US05004.  
PR 01-MAR-2000; 2000MO-US05601.  
PR 20-MAR-2000; 2000MO-US07377.  
PR 21-MAR-2000; 2000MO-US07377.  
PR 30-MAR-2000; 2000MO-US08439.  
PR 17-MAY-2000; 2000MO-US13705.  
PR 22-MAY-2000; 2000MO-US14042.  
PR 30-MAY-2000; 2000MO-US14941.  
PR 02-JUN-2000; 2000MO-US15264.  
PR 10-NOV-2000; 2000MO-US30873.

XX (GETH ) GENEENTECH INC.  
 PA Baker KP, Beresini M, DeForge L, Desnoyers L, Filyaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WT, Zhang Z;  
 XX WPI: 2001-408281/43.  
 DR P-PSDB: AAU12258.  
 XX  
 XX Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing  
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical  
 PS Claim 3: Fig 173: 813pp: English.  
 PS  
 XX  
 CC AAS21244-AAS21518 encode for novel human secretory and transmembrane  
 CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, to modulate biological activities of cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from  
 CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC of T-lymphocytes, the release of a cytokine from peripheral blood  
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIa. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding  
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy.  
 XX  
 SO Sequence 1210 BP: 261 A: 316 C: 379 G: 254 T: 0 other:  
 Query Match 93.0%; Score 749.4; DB 22; Length 1210;  
 Best Local Similarity 98.9%; Pred. No. 2.2e-179;  
 Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

DB 870 GCGCGGAATGAGAAATCCAGTGAACCTGACCTACAGAGCATCCAGATCAGATATT 929  
 OY 491 AGCGGACCCCTCAAGTTTCGTGTGAGACCATTTGTGAGAAATAGAGATGAACTCAT 550  
 DB 930 AGCGGACCCCTCAAGTTTCGTGTGAGACCATTTGTGAGAAATAGAGATGAACTCAT 989  
 OY 551 GAATTCCTTTCCCGAGAGAGCTACATATTTAAAGACAACTTTGCATTAAGCAACAT 610  
 DB 990 GAATTCCTTTCCCGAGAGAGCTACATATTTAAAGACAACTTTGCATTAAGCAACAT 1049  
 OY 611 CTTTGAGACCATCCCTGACATATTCGATGAGAGCTGATGAACCACTGAGACGCCAC 670  
 DB 1050 CTTTGAGACCATCCCTGACATATTCGATGAGAGCTGATGAACCACTGAGACGCCAC 1109  
 OY 671 ACTGGCTTGATGATGATCACCAGGAGGAGGAAATGTTGCAATGCTTTATATATAT 730  
 DB 1110 ACTGGCTTGATGATGATCACCAGGAGGAGGAAATGTTGCAATGCTTTATATATAT 1169  
 OY 731 GTTTTACTGAAATTAACGAAAAAATATGAACCAAAAGT 771  
 DB 1170 GTTTTACTGAAATTAACGAAAAAATATGAACCAAAAGT 1210  
 RESULT 9  
 ID AAX97884 standard; cDNA: 714 BP.  
 XX AAX97884;  
 XX 23-SEP-1999 (first entry)  
 DE Human secreted protein encoding cDNA #72.  
 KW Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;  
 KW diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
 OS Homo sapiens.  
 XX  
 XX W09925825-A2.  
 XX  
 PD 27-MAY-1999.  
 XX  
 PE 13-NOV-1998; 98WO-1B01862.  
 XX  
 PR 04-SEP-1998; 98US-0099273.  
 PR 13-NOV-1997; 97US-0066677.  
 PR 17-DEC-1997; 97US-0066957.  
 PR 09-FEB-1998; 98US-0074121.  
 PR 13-APR-1998; 98US-0081563.  
 PR 10-AUG-1998; 98US-0096116.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Bougueleret L, Duclert A, Dumas MLine Edwards J;  
 DR WPI: 1999-347472/29.  
 XX P-PSDB: AAY36200.  
 PT  
 PT Extended cDNAs encoding secreted proteins  
 PS Claim 1: Page 272; 307pp: English.  
 XX  
 CC AAY97813-X97906 represent extended cDNA's which encode novel human  
 CC secreted proteins (see AAY36129-Y36222) and which have cytosolic,  
 CC thrombotic and osteopathic activity. The extended cDNAs can be used to  
 CC express secreted proteins or parts of them or to obtain antibodies  
 CC capable of binding to the secreted proteins. They may also be used in  
 CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC uses also include design of expression vectors and secretion vectors.  
 XX  
 SO Sequence 714 BP: 203 A: 158 C: 202 G: 151 T: 0 other:

Query Match 87.1%; Score 702.4; DB 20; Length 714;  
 Best Local Similarity 98.9%; Pred. No. 1.3e-167;  
 Matches 706; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 72 ACCAAGGTGGAGCAACCCGTTACCTTAARATGAAGAGCGGGTGGCGCCCTCG 131  
 1 ACCAGAGGTGGAGCGACCCATTACCTTAAGATGAAGAGCGGGTGGCGCCCTCG 60  
 DB 132 TTCTGGGGGCGCTGTGGGAGACCGCTGGGCTGGAGGAGAGATCTCCACTGGAG 191  
 61 TTCTGGGGGCGCTGTGGGAGACCGCTGGGCTGGAGGAGAGATCTCCACTGGAG 120  
 QY 192 CATGCAAGGCTGTGGTGAATGAATGAATGGAAATTCGCCAGGTGAGCCCAAGAGA 251  
 121 CATGCAAGGCTGTGGTGAATGAATGAATGGAAATTCGCCAGGTGAGCCCAAGAGA 180  
 QY 252 CCATTCAATGGATCTTCCGATCAATCCAGATGGCCAGTCAGTGGTGGAGTGC 311  
 181 CCATTCAATGGATCTTCCGATCAATCCAGATGGCCAGTCAGTGGTGGAGTGC 240  
 QY 312 CTTATGCCGCTCAGAGGCCCACTCAGAGCTGCTGGAGGAGATATGTGACCGGATGA 371  
 241 CTTATGCCGCTCAGAGGCCCACTCAGAGCTGCTGGAGGAGATATGTGACCGGATGA 300  
 QY 372 ACGAGTATGGGAGACAGATTGATCTTCCACCACCAATCGAAGACTAGCTGTAGTG 431  
 301 ACGAGTATGGGAGACAGATTGATCTTCCACCACCAATCGAAGACTAGCTGTAGTG 360  
 QY 432 GCCGGAATGGAGAAATTCATGAACTGAACTGAAAGGCAATCCGAATCGACTCAGATTA 491  
 361 GCCGGAATGGAGAAATTCATGAACTGAACTGAAAGGCAATCCGAATCGACTCAGATTA 420  
 QY 492 GCGGCAACCTCAAGTTGGGTGAGAGCATTTGGAGGAAATACGAGATGAATCATTTG 551  
 421 GCGGCAACCTCAAGTTGGGTGAGAGCATTTGGAGGAAATACGAGATGAATCATTTG 480  
 QY 552 AATTCTTTTCCGAGAGCGTGACAAATGTTAAAGACAACTTTGGACATAAGCGACGATC 611  
 481 AATTCTTTTCCGAGAGCGTGACAAATGTTAAAGACAACTTTGGACATAAGCGACGATC 540  
 QY 612 TTGTGACCATGCGCTGCATATATGATGATGATGATGATGATGATGATGATGATGATG 671  
 541 TTGTGACCATGCGCTGCATATATGATGATGATGATGATGATGATGATGATGATGATG 600  
 QY 672 CTGGCTGTGATGATCAACCCCAAGAGGAGAAATGGTGGCAATGCTTTATATATATG 731  
 601 CTGGCTGTGATGATCAACCCCAAGAGGAGAAATGGTGGCAATGCTTTATATATATG 660  
 QY 732 TTTTACTGAAATTAATGAAAAAATATGAAACCAAAAGTAAAAAATAAAAA 785  
 661 TTTTACTGAAATTAATGAAAAAATATGAAACCAAAAGTAAAAAATAAAAA 714  
 DB

RESULT 10  
 AAX97837  
 ID AAX97837 standard; cDNA; 832 BP.  
 XX AAX97837;  
 AC  
 XX  
 DT 23-SEP-1999 (first entry)  
 XX  
 DE Human secreted protein encoding cDNA #25.  
 XX  
 KW Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;  
 XX diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
 OS Homo sapiens.  
 XX  
 PN W09925825-A2.  
 XX  
 PD 27-MAY-1999.  
 XX  
 PF 13-NOV-1998; 98WO-1B01862.

XX 04-SEP-1998; 98US-0099273.  
 PR 13-NOV-1997; 97US-0066677.  
 PR 17-DEC-1997; 97US-0069957.  
 PR 09-FEB-1998; 98US-0074121.  
 PR 13-APR-1998; 98US-0081563.  
 PR 10-AUG-1998; 98US-0096116.  
 XX  
 PA (GEST) GENSET.  
 XX  
 PI Bouquelier L, Duclert A, Dumas Milne Edwards J;  
 XX  
 DR WPI; 1999-347472/29.  
 DR P-PSDB; AAY36153.  
 XX  
 PT Extended cDNAs encoding secreted proteins  
 PS Example 28; Page 194-195; 307pp; English.  
 XX  
 CC AAX97813-X97906 represent extended cDNA's which encode novel human  
 CC secreted proteins (see AAY36129-Y36222) and which have cytosolic,  
 CC thrombotic and osteopathic activity. The extended cDNAs can be used to  
 CC express secreted proteins or parts of them or to obtain antibodies  
 CC capable of binding to the secreted proteins. They may also be used in  
 CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC Uses also include design of expression vectors and secretion vectors.  
 XX

Sequence 832 BP; 224 A; 194 C; 243 G; 168 T; 3 other;

Query Match 85.4%; Score 688.4; DB 20; Length 832;  
 Best Local Similarity 97.6%; Pred. No. 4.9e-164;  
 Matches 761; Conservative 1; Mismatches 12; Indels 6; Gaps 6;

QY 11 CTGGGGCCAAAGGAGAAAGTCCAGGCTGCGCCAGGCTGGGCGGAGGCGGCGCCCTGG 70  
 55 CTGGGGCCAAAGTGAAGATCCAGGCTGCTCCAGGCTGGGCGGAGGCGGCGCCCTGG 114  
 QY 71 GACCAAGGTGGAGCAACCCCGTTACCTTAARATGAAGAGCGTGGGTTGGTGGCCCTG 130  
 115 GACCAAGGTGGAGCAACCCCGTTACCTTAARATGAAGAGCGTGGGTTGGTGGCCCTG 174  
 QY 131 CTTCTGGGGCGCTGCTGGGAGCCGCTGGGCTGGAGGAGCCAGATCTCCACTGTGGA 190  
 175 CTTCTGGGGCGCTGCTGGGAGCCGCTGGGCTGGAGGAGCCAGATCTCCACTGTGGA 234  
 QY 191 GCATGCAAGGCTGTGGTGTG-ACATGAAATGGAAATTCGCCAGGTGAGCCCAAGAA 249  
 235 GCATGCAAGGCTGTGGTGTG-ACATGAAATGGAAATTCGCCAGGTGAGCCCAAGAA 294  
 QY 250 GACCATTCAGATGGATCTTTCCGATCAATCCAGATGGCAGCCAGTCAAGTGGTAGGT 309  
 295 GACCATTCAGATGGATCTTTCCGATCAATCCAGATGGCAGCCAGTCAAGTGGTAGGT 354  
 QY 310 GCCTTATGCCCGCTCAGAGGCCCACTCAGAGCTGCTGGAGAGATATGACCGGAT 369  
 355 GCCTTATGCCCGCTCAGAGGCCCACTCAGAGCTGCTGGAGAGATATGACCGGAT 414  
 QY 370 GAAGAGATATGGGAGACAGATTGATCTTCCACCACCAATCGCAAGATAGCTAGTACT 429  
 415 GAAGAGATATGGGAGACAGATTGATCTTCCACCACCAATCGCAAGATAGCTAGTACT 474  
 QY 430 GGGCCGAATGGAGAAATCCAGTGAACCTCAACAGAGCATCCGATCACTCAGATAT 489  
 475 GGGCCGAATGGAGAAATCCAGTGAACCTCAACAGAGCATCCGATCACTCAGATAT 534  
 QY 490 TACGGCAGCCCTCAAG-TTTGCGTGTGAGAGCAATTGTGAGAGAAATACGAGATGAATCA 548  
 535 TACGGCAGCCCTCAAGBTGTGCGTGTGAGAGCAATTGTGAGAGAAATACGAGATGAATCA 594  
 QY 549 TTGAATTTCTTTCCGAGAGAGGTGACAAATGTTAAAGCAAACTTTCAGTAAACGAACG 608  
 595 TTGAATTTCTTTCCGAGAGAGGTGACAAATGTTAAAGCAAACTTTCAGTAAACGAACG 654  
 DB

0Y 609 ATCTTTGACCATGCCCCGACATATC-GCATGATGACCTTGAACCACTGAGACGCC 667  
DB 655 ATCTTTGACCATGCCCCGACATATGCGATGATGACCTTGAACCACTGAGACGCC 714  
0Y 668 CACACTGGCTTGATGATCACCAGG-AGGGAAATGATGGCAATGCTTTATATA 726  
DB 715 CACACTGGCTTGATGATCACCAGGAAAGGAAATGATGGCAATGCTTTATATA 774  
0Y 727 TTAATCTTTTACTGAAATTAATGAAAAATGAAAAAGTAAAAA 786  
DB 775 TTAATCTTTTACTGAAATTAATG-AAAAATGAAAAAGTAAAAA 832

RESULT 11  
ABK36007  
ID ABK36007 standard: cDNA: 657 BP.  
XX  
AC ABK36007;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE cDNA sequence #398 encoding novel human secreted protein.  
XX  
KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;  
KW Immune deficiency disorder; blood disorder; inflammatory disorder;  
KW Infectious disorder; allergic condition; neurodegenerative disorder;  
KW Liver fibrosis; coagulation disorder; gene therapy; antileukoblastic;  
KW tumour; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200177289-A2.  
XX  
PD 18-OCT-2001.  
XX  
PF 29-MAR-2001; 2001WO-US10232.  
XX  
PR 06-APR-2000; 2000US-195605P.  
XX  
PA (GENY) GENETICS INST INC.  
XX  
PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
PI Meberg D, Treacy M, Agostino MJ, Bowman MR, Spaulding V, Wong GG;  
PI Clark HF, Fechtel K, Howes SH, Resnick RJ, Gulikota K, Graham JR;  
XX  
DR WPI: 2002-179322/23.  
XX  
PT Six hundred and twenty three polynucleotides derived from a variety of  
PT human tissue sources which encode secreted proteins, useful for  
PT treating immune deficiencies and disorders such as autoimmune disorders  
PT  
XX  
PS Claim 1; Page 297; 393pp; English.  
XX  
CC The present invention relates to the isolation of novel cDNA sequences  
CC which encode human secreted proteins. The cDNA sequences have been  
CC derived from a variety of human tissues. The invention also provides  
CC a method for producing proteins from these polynucleotide sequences.  
CC The proteins are useful for identifying compounds that modulate their  
CC activity and production. The sequences of the invention are  
CC useful for treating diseases such as hyperproliferative disorders  
CC (e.g. cancer), immune deficiency disorders (e.g. severe combined  
CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple  
CC sclerosis), blood disorders (e.g. thrombocytopenia), inflammatory  
CC disorders (e.g. arthritis), infectious disorders (e.g. hepatitis),  
CC allergic conditions (e.g. asthma), neurodegenerative disorders (e.g.  
CC Alzheimer's disease), liver fibrosis, coagulation disorders (e.g.  
CC haemophilia), and tumours. The polynucleotide sequences of the  
CC invention are also useful in gene therapy. ABK35610-ABK36232 represent  
CC the cDNA sequences of the invention that encode for novel human  
CC secreted proteins.  
XX  
SO Sequence 657 BP; 177 A; 149 C; 185 G; 146 T; 0 other;

Query Match 80.2%; Score 646.4; DB 24; Length 657;  
Best Local Similarity 99.1%; Pred. No. 1.8e-153;  
Matches 650; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

0Y 116 GGTGGCTGGCCCTGCTTGGGGGCGCTGCTGGGAACCGCTGGGCTGGAGAGCCAG 175  
DB 1 GGTGGCTGGCCCTGCTTGGGGGCGCTGCTGGGAACCGCTGGGCTGGAGAGCCAG 60  
0Y 176 GATCTCCACTGTGAGCATGAGGCGCTGTGTGATGAATGGAATTTGCCAG 235  
DB 61 GATCTCCACTGTGAGCATGAGGCGCTGTGTGATGAATGGAATTTGCCAG 120  
0Y 236 GTTGACCCCAAGAACACCTTCAGATGATGATCTTCCGATCAATCCAGATGAGCAG 295  
DB 121 GTTGACCCCAAGAACACCTTCAGATGATGATGATCTTCCGATCAATCCAGATGAGCAG 180  
0Y 296 TCAGTGTGGAGAGTGGCTTATGCCCCCTAGAGGCCCACTCAGACAGCTGCTGGAGAG 355  
DB 181 TCAGTGTGGAGAGTGGCTTATGCCCCCTAGAGGCCCACTCAGACAGCTGCTGGAGAG 240  
0Y 356 ATATGTGACCGGATGAGAGATGAGGGAACAGATTGATCTTCCACCATGCAAGAAC 415  
DB 241 ATATGTGACCGGATGAGAGATGAGGGAACAGATTGATCTTCCACCATGCAAGAAC 300  
0Y 416 TACGTACGTGTAGTGGGCGCGAATGAGAAATCCAGTGAACCTGACCTACAGCATCCGA 475  
DB 301 TACGTACGTGTAGTGGGCGCGAATGAGAAATCCAGTGAACCTGACCTACAGCATCCGA 360  
0Y 476 ATGCACTGATATATAGCCGACCCCTCAAGATTGGCTGTGAAGCATTTGGAGCAATTC 535  
DB 361 ATGCACTGATATATAGCCGACCCCTCAAGATTGGCTGTGAAGCATTTGGAGCAATTC 420  
0Y 536 GAGATGAACTGATTAATCTTTCCGAGAGGCTGACAAATGTTAAAGCAAACTTTCG 595  
DB 421 GAGATGAACTGATTAATCTTTCCGAGAGGCTGACAAATGTTAAAGCAAACTTTCG 480  
0Y 596 ACTAAGCGAAGATCTTTGTGACCATGCGCTGACACATATGCGATGAGTGAAC 655  
DB 481 ACTAAGCGAAGATCTTTGTGACCATGCGCTGACACATATGCGATGAGTGAAC 540  
0Y 656 ACTGAGAGCGCCACACAGCGCTTGATGATGATGATGATGATGATGATGATGATGATG 715  
DB 541 ACTGAGAGCGCCACACAGCGCTTGATGATGATGATGATGATGATGATGATGATGATG 600  
0Y 716 CCTTTATATATATGTTTACTGAAATTAATGAAAAATATGAAACCAAAAGT 771  
DB 601 CCTTTATATATATGTTTACTGAAATTAATGAAAAATATGAAACCAAAAGT 656

RESULT 12  
AAK06970  
ID AAK06970 standard: cDNA: 1069 BP.  
XX  
AC AAK06970;  
XX  
DT 10-MAY-1999 (first entry)  
XX  
DE Mouse secretory peptide-9 (Zs1g9) orthologue cDNA.  
XX  
KW Secretory peptide-9; Zs1g9; orthologue; mouse; tumour marker;  
KW cancer; therapy; diagnosis; growth enhancer; ss.  
XX  
OS Mus sp.  
XX  
FH Key 358..906  
FH CDS 358..906  
FH s1g\_peptide 358..417  
FH /\*tag- b  
FT mat\_peptide 418..903  
FT /\*tag- c  
XX

```

XX MN09901554-A1.
PD 14-JAN-1999.
XX PF 02-JUL-1998; 98WO-US13859.
XX 17-JUN-1998; 98US-0099005.
PR 03-JUL-1997; 97US-0051704.
PR 03-JUL-1997; 97US-0886088.
PR 19-MAY-1998; 98US-0081338.
PR 19-MAY-1998; 98US-0085983.
PR 17-JUN-1998; 98US-0089899.
XX (ZYMO ) ZYMOGENETICS INC.
PA Jaspers SR, Jelinek LJ, Sheppard PO, Whitmore TE;
PI WPI: 1999-106055/09.
DR P-PSDB: AAM88476.
XX New mammalian secretory peptide-9 (Zs1g9) - used as a growth
PT enhancer for placenta, liver and heart, and as an indicator of
PT cancer
XX Claim 2; Page 75-77; 85pp; English.
PS This cDNA clone encodes novel mouse secretory peptide-9, or Zs1g9
XX (see AAM88476), an orthologue of novel human Zs1g9 (see AAM88465).
CC Human Zs1g9 is overexpressed in a number of tumours including
CC brain, liver, lung, oesophageal, stomach, colon, rectal, thyroid
CC and lymphoma tumors, and thus can be used as an indicator for
CC cancer. The invention provides polynucleotides (see AA06986-70)
CC encoding Zs1g9 polypeptides (see AAM88465-77) including mature
CC polypeptides, other processed forms, variants and the mouse
CC orthologues. The Zs1g9 gene, or probes derived from it, can be
CC used to determine if Zs1g9 is present on chromosome 10, and if a
CC mutation has occurred. Antibodies raised against Zs1g9 can be
CC used as diagnostic agents to determine the presence of Zs1g9, and
CC thus the presence of cancer. They can also be labelled with
CC radiolabels or fused with toxins and used to treat tumors
CC which overexpress Zs1g9. Antisense nucleotides derived from Zs1g9
CC cDNA can also be used to inhibit the growth of tumor cells. Zs1g9
CC proteins can be used to enhance the growth or development of the
CC placenta, heart or liver.
XX Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other;
SQ Query Match 74.3%; Score 599; DB 20; Length 1069;
Best Local Similarity 87.7%; Pred. No. 2,1e-141;
Matches 690; Conservative 1; Mismatches 86; Indels 10; Gaps
OY 18 CAAATGGAAGTCACAGCGGTCTGCAGGCCCTTGGCCACGCGGCGCCTTGAGCCAAA 77
Db 272 CAAATGAGAGTCGCGGCTCTTCCAGAGCCCTGGCCACGCGGCGCCTTGAGACAGA 331
OY 78 GGTGAGCAACCCCGTTACCCTTAARATGAAGAGCTGGGGCTGGCCCTTCCTTGG 137
Db 332 GTTGAGAGGACCCCTGTATACATAAAGATGAAGAGCTGGGGTGGGACTTCTTGG 391
OY 138 GGGCGCTGTGGGAACCGCGCTGGGCTCGGAGAGCCAGATCTCCACTGTGAGCATGCA 197
Db 332 GGGCTCTGTGGGAAGTGGCTGGGCTCGAAGAGCCCAATCTACACTGTGAGATTGCA 451
OY 198 GGGCTCTGTGATGAATAGATGGGAATTTGCCAGGTGAGCCCAAGAAAGACATTTC 257
Db 452 GGGCTCTGTGATGAATTAAGATGGGAATTTGCCCGCTGACCCCAAGAACCATTC 511
OY 258 AGATGGGATCTTTCGGGATCAATCCAGATGGAGCCAGTCAGTGTGAGAGTGCCTTAG 317
Db 512 AGATGGGATCTTTCGGAATCAATCCAGATGGAGCCAGTCAGTGTGAGAGTACCTTAG 571
OY 318 CCCGCTCAGAGGCCCACTCACAGAGCTCTGTGAGGAGATATGTACCGGATGAAGAGT 377

```

DB	5712	CCGGCTCAGAGGCCACACCTCACAGAGTTGCTTGGAGAGGTGTGTGACCCAAATGAAGGACT	631
OY	378	ATGGGGAACAGATTATATCTCTTCACCCATCGCAAGAACTACTAGCTGTAGTGGCCGGA	437
Db	632	ACGGGGAACAGATTATATCTCTTCACCCATCGCAAGAACTACTAGCTGTAGTGGCCGGA	691
OY	438	ATGGAGAAATCCAGTAACCTGAGACCTACAAAGGATCCGAATTCAGATCATATTATTAAGCGCA	497
Db	692	ATGGAGAAATCCAGTAACCTGAGACCTACAAAGGATCCGAATTCAGATCATATTATTAAGCGCA	751
OY	438	CCCTCAAGTTTGCCTGTGAGACATTGTGGAGAAATACGAGGATGAAATCATTAATCT	557
Db	752	CCCTCAAGTTTGCCTGTGAGACATTGTGGAGAAATACGAGGATGAAATCATTAATCT	811
OY	558	TTTCCGAGAGGCTGACAACTGTTAAAGCAAACTTGCAGTAAGCGAACAGATCTTTGTG	617
Db	812	TCTCCAGAGAGGCTGACAAAGCTTTAAAGCAAACTTGCAGTAAGCGAACAGATCTTTGTG	871
OY	618	ACCAATCCCTGACATATTCGCAATGATGACTTGAACCACTGAGACGCCAC-----AC	672
Db	872	ACCAATCCCTGACATATTCGCAATGATGACTTGAATCACTGAGACGCCACCTACAC	931
OY	673	TGGCTTATGATGATACCCCGAGAGGGGAAATGGGCAATGCCCTTTATATATATGT	732
Db	932	AAACGTGATGAGAACACCCCGAGAGGGGAAATGGGCAATGCCAGCTTATGCTTT--TATATTACGT	989
OY	733	TTTTACTGAATAATTAATCTGAAAAA--TATGAACCCAAAGTAAAAAATTTTTTTTTT	789
Db	990	TTTTATGAAATGAAGACTGAAAAAATCTGTAAACCGAAGTAAAAAATTTTTTTTTT	1049
OY	790	AGAGAGA 796	
Db	1050	AAAAAAA 1056	
RESULT 13			
ID	AAZ08294		
ID	AAZ08294	standard; DNA; 1069 BP.	
XX	AAZ08294;		
AC			
XX			
DT	07-FEB-2000	(first entry)	
DE			
XX			
XX			
KW	Secretory protein-9; mouse zslg9; ortholog; overexpression; antagonist;		
KW	antibody; antisense nucleotide; tumour; treatment; receptor;		
KW	radio-label; polypeptide toxin; down-regulation; diagnostic;		
KW	therapeutic; probe; cancer; brain; liver; detection; stomach;		
KW	lymphoma; ds.		
OS	Mus musculus.		
XX			
EH	Key	Location/Qualifiers	
FT	CDS	358..906	
FT		/*tag= a	
FT		/product= "Mouse zslg9 secretory protein"	
FT		/note= "Overexpressed in tumours"	
FT	sig_peptide	358..417	
FT		/*tag= b	
FT	mat_peptide	418..903	
FT		/*tag= c	
XX		/label= Mature_zslg9_protein_variant-4	
PN	W09960405-A1.		
XX			
PD	25-NOV-1999.		
XX			
PF	19-MAY-1999;	99MO-US1107.	
XX			
PR	19-MAY-1998;	98US-0081183.	
XX			
PA	(ZYMO ) ZYMOGENETICS INC.		

```

XX Moore BE, Taft DW;
PI WPI: 2000-039447/03.
DR P-PSDB: AAY15136.
XX
XX Detecting tumors using antibodies, antagonists and antisense
XX nucleotides to secretory protein-9 (Zs19)
XX
XX Disclosure; Page 37-38; 45pp; English.
XX
XX The present DNA sequence is an ortholog encoding the secretory protein-9,
XX Zs19 derived from mouse. It is overexpressed in tumors. Antagonists,
XX antibodies and antisense nucleotides to Zs19 are useful for detecting
XX and treating tumors. The antagonist may be an antibody or receptor to
XX Zs19 and it may be radio-labelled or fused to a polypeptide toxin. It
XX can be used for down regulating the overexpression of Zs19. The gene
XX sequence can be used as nucleic acid probes to detect RNA encoding Zs19.
XX The Zs19 sequence facilitates improved diagnostic and therapeutic
XX techniques for detecting and treating cancers, especially of the brain,
XX liver, stomach, lymphoma etc., at an early stage.
XX
SQ Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other:
Query Match 74.3%; Score 599; DB 21; Length 1069;
Best Local Similarity 87.7%; Pred. No. 2.1e-141;
Matches 690; Conservative 1; Mismatches 86; Indels 10; Gaps 3;
OY 18 CAAAGTGAAGTCCACGCGTCTGCCAGCCTTGCGGCACGCGCGCGCCCTGGACCATA 77
DB 272 CAAAGTGAAGTCCACGCGTCTGCCAGCCTTGCGGCACGCGCGCGCCCTGGACCATA 331
OY 78 GGTGAGACACCCCGTTACCTTAATGAAGAGCTGGGGTGGCTGCGCTGCTTCTG 137
DB 332 GGTGAGACACCCCGTTACCTTAATGAAGAGCTGGGGTGGCTGCGCTGCTTCTG 391
OY 138 GGGGCCCTGTGGGAAACCGCTGGCGGAGGAGGAGATCTCCACTGTGAGCATGCA 197
DB 392 GGGTCTGTGGGAAACCGCTGGCGGAGGAGGAGATCTCCACTGTGAGCATGCA 451
OY 198 GGGCTCTGTGATGAATAGATGGGAATGGCCAGGTGAGCCCAAGAACCATTC 257
DB 452 GGGCTCTGTGATGAATAGATGGGAATGGCCAGGTGAGCCCAAGAACCATTC 511
OY 258 AATGGATCTTCCGATCAATCCAGATGCAAGCCAGTGTGAGAGTCTTATG 317
DB 512 AATGGATCTTCCGATCAATCCAGATGCAAGCCAGTGTGAGAGTCTTATG 571
OY 318 CCCGCTCAGAGGCCACCTCAGAGAGCTGTGAGAGATGTGACCGGATGAAGAGT 377
DB 572 CCCGCTCAGAGGCCACCTCAGAGAGCTGTGAGAGATGTGACCGGATGAAGAGT 631
OY 378 ATGGGAAACAGATTGATCTCTCCACCATCGCAAGAACTAGCTGTAGTGGCCGA 437
DB 632 ATGGGAAACAGATTGATCTCTCCACCATCGCAAGAACTAGCTGTAGTGGCCGA 691
OY 438 ATGGAAATCCAGTGAATGACCTTCAAGCAAGCAATCCGAATCACTGATTAAGCGGA 497
DB 692 ATGGAAATCCAGTGAATGACCTTCAAGCAAGCAATCCGAATCACTGATTAAGCGGA 751
OY 498 CCCTCAAGTTGCGTGTGAGAGCATTTGTGAGGAATACAGAGATGAACATGGAATTCT 557
DB 752 CCCTCAAGTTGCGTGTGAGAGCATTTGTGAGGAATACAGAGATGAACATGGAATTCT 811
OY 558 TTTCCGAGAGGCTGCATGTTAAAGCAAACTTGCAGTAAGGACAGACATCTTGTG 617
DB 812 TTTCCGAGAGGCTGCATGTTAAAGCAAACTTGCAGTAAGGACAGACATCTTGTG 871
OY 618 ACCATGCTCTGCATATGCAATGATGATGATGAACCACTGAGAGCCAC-----AC 672
DB 872 ACCATGCTCTGCATATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 931
OY 673 TGGCTTGAATGATCAACCCCGAGAGGGGAAATGATGATGATGATGATGATGATGATGAT 732

```

```

DB 932 AAACGTGATGGAACACCCCGAGAGGGAAGATGCGACGATTCCTTT--TATATTACGT 989
OY 733 TTTTACTGAATTAATGAGAAAAA---TATGAACCAAAAGTAAAAAAGG 789
DB 990 TTTTATGAAATGAACTGAAAAAACTTTGAAACCGAAGTAAAAAAGGAAAAA 1049
OY 790 AGAGAGA 796
DB 1050 AAAAAA 1056

RESULT 14
AAZ38326
ID AAZ38326 standard; cDNA; 546 BP.
XX
XX AAZ38326;
AC
XX 09-FEB-2000 (first entry)
DT
XX
XX Human transmembrane protein cDNA clone HP10390 coding sequence.
DE
XX
XX HP10390; transmembrane domain; stomach cancer cell; antibody;
KW assay reagent; diagnostic marker; primer; probe; antisense; gene therapy;
KW agonist; antagonist; ligand; therapeutic; ds.
XX
XX Homo sapiens.
XX
XX Key location/Qualifiers
FH CDS 1..540
FT
FT domain in the N-terminus. The cDNA was isolated from a human stomach
FT /product= "Human transmembrane protein HP10390"
FT /note= "No stop codon given in the specification"
XX
XX W09955862-A2.
XX
XX 04-NOV-1999.
XX
XX 27-APR-1999; 99WO-JP02226.
XX
XX 28-APR-1998; 98JP-0119395.
XX
XX (SAGA ) SAGAMI CHEM RES CENT.
XX (PROT-) PROTEGENE INC.
XX
XX Kato S, Kimura T;
XX
XX WPI: 2000-023358/02.
XX P-PSDB: AAY52391.
XX
XX Human proteins with transmembrane domains, involved in control of cell
XX proliferation and differentiation, useful for treating e.g. cancer or
XX inflammation
XX
XX Claim 3; Page 88; 114pp; English.
XX
XX This sequence represents the coding sequence of human cDNA clone
XX HP10390 which encodes a 20 kD protein with one putative transmembrane
XX domain in the N-terminus. The cDNA was isolated from a human stomach
XX cancer cell line cDNA library. The protein has no homology with any
XX known protein. The protein may be used to raise specific antibodies, as
XX assay reagents, as diagnostic tissue markers, for the isolation of
XX cognate receptors, ligands and binding proteins, and as biologically
XX active agents. Nucleotides encoding the protein may be used as primers
XX and probes or antisense molecules, and in gene therapy. Cells transfected
XX with these nucleotides may be used to screen for agonists and antagonists
XX which are potentially useful therapeutically.
XX
SQ Sequence 546 BP; 139 A; 126 C; 165 G; 116 T; 0 other:
Query Match 67.7%; Score 546; DB 21; Length 546;
Best Local Similarity 100.0%; Pred. No. 3.9e-128;
Matches 546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 104 ATGAAAGCTGGGCTGGCTGGCCCTGCTTGGGGGCCCTGCTGGGAACCGCTGGGCT 163  
DB 1 ATGAAAGCTGGGCTGGGCTGGCCCTGCTTGGGGGCCCTGCTGGGAACCGCTGGGCT 60  
QY 164 CGAGAGACCCAGATCTCCACTGTGAGACATGAGGCGCTGTGTGATGATGATGAGT 223  
DB 61 CGAGAGACCCAGATCTCCACTGTGAGACATGAGGCGCTGTGTGATGATGATGAGT 120  
QY 224 GAAATGGCCAGGTGGACCCCAAGAACCATTCAGATGGGATCTTCCGATCAATCCA 283  
DB 121 GAAATGGCCAGGTGGACCCCAAGAACCATTCAGATGGGATCTTCCGATCAATCCA 180  
QY 284 GATGAGCCAGCTAGTGTGTGAGTGGCTTATGCCCGCTCAGAGGCCACCTCAGAGAG 343  
DB 181 GATGAGCCAGCTAGTGTGTGAGTGGCTTATGCCCGCTCAGAGGCCACCTCAGAGAG 240  
QY 344 CTGCTGGAGAGATATGTGACCCGATGAGAGATATGGGGAACAGATTCCTCCACC 403  
DB 241 CTGCTGGAGAGATATGTGACCCGATGAGAGATATGGGGAACAGATTCCTCCACC 300  
QY 404 CATCGAAGAACTACGTAGTGTGAGGCGCGAATGAGAACTCAGTGAATGAGACTTA 463  
DB 301 CATCGAAGAACTACGTAGTGTGAGGCGCGAATGAGAACTCAGTGAATGAGACTTA 360  
QY 464 CAAGGATCCGAATCGACTGAGATATTTAGCGGACCCCTCAAGTTGGCTGAGAGCAT 523  
DB 361 CAAGGATCCGAATCGACTGAGATATTTAGCGGACCCCTCAAGTTGGCTGAGAGCAT 420  
QY 524 GTGAGGAATACGAGATGAGTCAATTCATTCCTTCCGAGAGGCTGACATGTTAA 583  
DB 421 GTGAGGAATACGAGATGAGTCAATTCATTCCTTCCGAGAGGCTGACATGTTAA 480  
QY 584 GACAACTTTCGATGAGGAGCAAGATCTTGTGACCATCCCTGACATATGATGAT 643  
DB 481 GACAACTTTCGATGAGGAGCAAGATCTTGTGACCATCCCTGACATATGATGAT 540  
QY 644 GAGCTA 649  
DB 541 GAGCTA 546

RESULT 15  
AAA43525  
ID AAA43525 standard; cDNA; 592 BP.  
XX  
AC AAA43525;  
XX  
DT 21-AUG-2000 (first entry)  
XX  
DE Mouse secreted expressed sequence tag SEQ ID NO:100.  
XX  
XX Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;  
XX expressed sequence tag; EST; probe; chemotactic; proliferative;  
XX immunomodulatory; haematopoietic; chemokine; analgesic; haemostatic;  
XX thrombolytic; antiinflammatory; cytosolic; antibacterial; antifungal;  
XX antiviral; antidiabetic; antiasthmatic; vulnery; antiparkinsonian;  
XX anticancer; osteoprotective; neuroprotective; nootropic; antipsoriatic;  
XX cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
XX vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
XX insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
XX lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
XX central nervous system disorder; Alzheimer's disease; stroke;  
XX Parkinson's disease; Huntington's disease; coagulation disorder;  
XX haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
XX tumour; infection; depression; psoriasis; ss.  
OS Mus musculus.  
XX  
XX MO200021991-A1.  
XX  
XX PD 20-APR-2000.

PF 15-OCT-1999; 99WO-US24206.  
XX  
PR 15-OCT-1998; 98US-0104436.  
XX  
PA (GENE) GENETICS INST INC.  
PI Jacobs K, McCoy JM, Lavallee ER, Collins-Racie LA, Evans C;  
PI Merberg D, Treacy M, Bowman MR;  
XX  
XX MPI: 2000-317938/27.  
PT Isolated polynucleotides, and encoded proteins, comprising secreted  
PT expressed sequence tags (SESTs), useful for treating various disorders  
PT such as autoimmune, infectious, and central nervous system disorders -  
XX  
XX Claim 1: Page 228-229; 803pp; English.  
XX  
CC AAA43426 to AAA45925 represent specifically claimed secreted expressed  
CC sequence tags (SESTs), isolated from human, mouse, chicken and rat  
CC tissue sources. The SESTs can have a range of activities depending on  
CC the tissues they were isolated from. The activities include:  
CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;  
CC cytosolic; antibacterial; antifungal; antiviral; antidiabetic;  
CC antiasthmatic; vulnery; anticonvulsant; antiparkinsonian;  
CC neurotropic; antiparkinsonian; antipsoriatic; cerebroprotective;  
CC anticonvulsant; and antidepressant. The SESTs can be used for gene  
CC therapy and in vaccines. The SESTs are useful as probes for the  
CC identification and isolation of full-length cDNAs and genomic DNA  
CC molecules which correspond to the SESTs. Proteins encoded by the SESTs  
CC are useful in assays for determining biological activity and raising  
CC antibodies. They may be useful for treatment of autoimmune disorders  
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
CC osteoporosis, osteoarthritis, central nervous system disorders  
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
CC disease), tumours, bacterial, fungal or viral infections, depression and  
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
CC in the exemplification of the present invention.  
XX  
SQ Sequence 592 BP; 159 A; 139 C; 167 G; 123 T; 4 other;  
Query Match 54.5%; Score 439.2; DB 21; Length 592;  
Best Local Similarity 89.8%; Pred. No. 3; 8e-101;  
Matches 501; Conservative 1; Mismatches 53; Indels 3; Gaps 3;  
QY 80 TGGAGCAACCCGTTACCCAAATGAAAGCGTGGGCTGGCTGGG-CCTGCTTGGG 138  
DB 20 TGGAGCAACCCGTTACCCAAATGAAAGCGTGGGCTGGGCTGGGCTTGGG 79  
QY 139 GGCCTGCTGGGAAACCGCTGGGCTGGGAGAGACGATCTCCACTGTGAGCATGAC 198  
DB 80 GGCCTGCTGGGAAACCGCTGGGAGATGCAAGAGACCAANATCTACACTGTGAGCATG 139  
QY 199 GGCCTGCTGGGATGATGATGAAATGGGAAATTTGCCAGGTGAGACCCCAAGACCATCA 258  
DB 140 GGCCTGCTGGGATGATGATGAAATGGGAAATTTGCCAGGTGAGACCCCAAGACCATCA 199  
QY 259 GATGAGATCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 318  
DB 200 GATGAGATCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 259  
QY 319 CCGCTCAGAGCCACCTCAGAGATGCTGAGAGATATGATGATGATGATGATGATGAT 378  
DB 260 CCGCTCAGAGCCACCTCAGAGATGCTGAGAGATATGATGATGATGATGATGATGAT 319  
QY 379 TGGAGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 438  
DB 320 -CGGAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 377  
QY 439 TGGAGATCCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 498

Db 378 TGGAGATCCAGTGAACCTAGACTTACAGGGCATCCGAATTGACTCAGATATTCANCGGCAC 437  
OY 499 CCTCAAGTTTGGCTGTGAGAGCATTTGGAGAGATACGAGATGAATCATTGAAATTCCTT 558  
|||||  
Db 438 CCTCAAGTTTGGCTGTGAGAGCATTTGGAGAGATACGAGATGAATCATTGAAATTCCTT 497  
OY 559 TTCCGAGAGGCTGACATATTTAAAGACAACCTTTGACAGTAAAGCGACAGATCTTGTGA 618  
|||||  
Db 498 CTCAGAGAGGCTGACATATTTAAAGACAACCTTTGACAGTAAAGCGACAGATCTTGTGA 557  
OY 619 CCATGCCCTGCACATATTC 636  
|||||  
Db 558 CCATGCCCTGCACAGATTC 575

Search completed: March 30, 2003, 18:47:31  
Job time : 197.291 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 18:39:57 ; Search time 44.7061 Seconds  
(without alignments)  
5529.028 Million cell updates/sec

Title: US-10-082-502-16

Perfect score: 806  
Sequence: 1 cggcccaagcctggggccaa.....aaagagagagagaacta 806

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Database :

1: /cgn2\_6/ptodata/1/lna/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/lna/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/lna/5A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/lna/5B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/lna/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/lna/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	48	6.0	1864	4	US-09-149-476-130 Sequence 130, App1
2	44.6	5.5	1395	4	US-08-553-367A-1 Sequence 1, App1
3	44.6	5.5	1395	4	US-09-295-306-1 Sequence 1, App1
4	44.6	5.5	1395	4	US-09-734-719-1 Sequence 1, App1
5	44.6	5.5	7218	1	US-08-332-463-14 Sequence 14, App1
6	43.8	5.4	1813	5	PCT-US94-12883-3 Sequence 3, App1
7	43.8	5.4	1898	1	US-08-342-411A-1 Sequence 1, App1
8	43.2	5.4	2296	4	US-08-496-841C-137 Sequence 137, App1
9	42.2	5.2	3214	1	US-08-484-105-17 Sequence 17, App1
10	42.2	5.2	3214	1	US-08-484-105-17 Sequence 17, App1
11	42.2	5.2	3300	4	US-09-336-643A-82 Sequence 82, App1
12	41.6	5.2	629	4	US-09-385-982-389 Sequence 389, App1
13	41.6	5.2	2246	4	US-09-363-708-3 Sequence 3, App1
14	41.6	5.2	7859	1	US-07-854-596B-4 Sequence 4, App1
15	41.6	5.2	7859	2	US-08-450-905B-15 Sequence 15, App1
16	41.6	5.2	7859	3	US-07-982-759F-15 Sequence 15, App1
17	41.4	5.1	1075	4	US-08-400-006B-6 Sequence 6, App1
18	41.4	5.1	1212	4	US-09-182-145-34 Sequence 34, App1
19	41.4	5.1	1212	4	US-09-182-145-35 Sequence 35, App1
20	41.4	5.1	1582	3	US-08-545-196B-10 Sequence 10, App1
21	41.4	5.1	1582	3	US-08-545-196B-12 Sequence 12, App1
22	41.4	5.1	1193	4	US-09-372-422A-23 Sequence 23, App1
23	41.4	5.1	1737	1	US-08-202-056-4 Sequence 4, App1
24	41.4	5.1	1737	1	US-08-076-093A-3 Sequence 3, App1
25	41.4	5.1	1737	1	US-08-701-265-3 Sequence 3, App1
26	41.4	5.1	1737	1	US-08-284-586-3 Sequence 3, App1
27	41.4	5.1	1737	2	US-08-805-478-3 Sequence 3, App1

28	41	5.1	1737	2	US-08-802-627A-3 Sequence 3, App1
29	41	5.1	1737	2	US-08-801-238-3 Sequence 3, App1
30	41	5.1	1737	2	US-08-801-228-3 Sequence 3, App1
31	41	5.1	1737	3	US-09-104-296-3 Sequence 3, App1
32	41	5.1	1737	5	PCT-US94-06380-2 Sequence 2, App1
33	40.6	5.0	1117	4	US-09-247-373B-33 Sequence 33, App1
34	40.6	5.0	1639	2	US-08-737-524B-1 Sequence 1, App1
35	40.6	5.0	2186	4	US-09-360-545-66 Sequence 66, App1
36	40.4	5.0	1129	4	US-09-227-357-40 Sequence 40, App1
37	40.2	5.0	140	1	US-08-628-417-5 Sequence 5, App1
38	40.2	5.0	240	1	US-08-628-417-6 Sequence 6, App1
39	40.2	5.0	1485	4	US-09-372-422A-39 Sequence 39, App1
40	40.2	5.0	2205	3	US-08-888-077A-41 Sequence 41, App1
41	40.2	5.0	2301	4	US-09-332-191-8 Sequence 8, App1
42	40.2	5.0	2301	4	US-09-232-200-8 Sequence 8, App1
43	40.2	5.0	2301	4	US-09-232-197-8 Sequence 8, App1
44	40.2	5.0	2301	4	US-09-232-201-8 Sequence 8, App1
45	40.2	5.0	2710	4	US-09-232-200-44 Sequence 44, App1

## ALIGNMENTS

RESULT 1  
US-09-149-476-130  
Sequence 130, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23



EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 6.0%; Score 48; DB 4; Length 1864;  
Best Local Similarity 63.2%; Pred. No. 0.0012;  
Matches 72; Conservative 1; Mismatches 41; Indels 0; Gaps 0;

QY 690 CCGAGGAGGGAATGTCGCAATGCTTTATATATATGTTTACTGAAATTAAT 749  
DB 1747 CCGAGGAGGGAATGTTTATATATGTCATGATTAAGTTGTTTGAATAAKAAA 1806  
QY 750 GAAAAATATGAACCAAGTAAAAAAGAGAGAGAGAGAGA 803  
DB 1807 AA 1860

## RESULT 2

US-08-553-367A-1  
Sequence 1, Application US/08553367A  
Patent No. 5939539

GENERAL INFORMATION:  
APPLICANT: Theodor LANGE et al.  
TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
STREET: 2033 K Street, N.W., Suite 800  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/553,367A  
FILING DATE: No. 5939539ember 27, 1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Lee Cheng

REGISTRATION NUMBER: 40,949

REFERENCE/DOCKET NUMBER: 49/FD4, 5M2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-721-8200

TELEFAX: 202-721-8250

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1395 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Cucurbita maxima

IMMEDIATE SOURCE:

CLONE: Clone pB11

US-08-553-367A-1

Query Match 5.5%; Score 44.6; DB 2; Length 1395;  
Best Local Similarity 63.6%; Pred. No. 0.0091;  
Matches 68; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 692 CAGAGGAGGGAATGTCGCAATGCTTTATATATATGTTTACTGAAATTAATCGA 751  
DB 1289 CTGGAAGTGCATTAATAAACCCTTTAATACATGTTCTTACTTGGCTCAAAAAA 1348  
QY 752 AAAAAATATGAACCAAGTAAAAAAGAGAGAGAGA 798  
DB 1349 AA 1395

## RESULT 3

US-09-295-306-1  
Sequence 1, Application US/09295306  
Patent No. 6198021

GENERAL INFORMATION:

APPLICANT: Theodor LANGE et al.

TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.

STREET: 2033 K Street, N.W., Suite 800

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/295,306

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/553,367

FILING DATE: No. 6198021ember 27, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Lee Cheng

REGISTRATION NUMBER: 40,949

REFERENCE/DOCKET NUMBER: 49/DIV-FD4, 5M2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-721-8200

TELEFAX: 202-721-8250

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1395 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Cucurbita maxima

IMMEDIATE SOURCE:

CLONE: Clone pB11

US-09-295-306-1

Query Match 5.5%; Score 44.6; DB 4; Length 1395;  
Best Local Similarity 63.6%; Pred. No. 0.0091;  
Matches 68; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 692 CAGAGGAGGGAATGTCGCAATGCTTTATATATATGTTTACTGAAATTAATCGA 751  
DB 1289 CTGGAAGTGCATTAATAAACCCTTTAATACATGTTCTTACTTGGCTCAAAAAA 1348  
QY 752 AAAAAATATGAACCAAGTAAAAAAGAGAGAGAGA 798  
DB 1349 AA 1395

```

: RESULT 4
: US-09-734-719-1
: Sequence 1, Application US/09734719
: Patent No. 645675
: GENERAL INFORMATION:
: APPLICANT: Theodor LANGE et al.
: TITLE OF INVENTION: GA-20-OXIDASE GENE SEQUENCES (As Amended)
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
: STREET: 2033 K Street, N.W., Suite 800
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20006
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 Inch, 1.44 mb
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: Wordperfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/734,719
: FILING DATE: 13-Dec-2000
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/295,306
: FILING DATE: April 21, 1999
: APPLICATION NUMBER: 08/553,367
: FILING DATE: No. 645675ember 27, 1995
: APPLICATION NUMBER: PCT/EP94/01664
: FILING DATE: May 24, 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Lee Cheng
: REGISTRATION NUMBER: 40,949
: REFERENCE/DOCKET NUMBER: 2000-1678/LC/01784
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-721-8200
: TELEFAX: 202-721-8250
: TELEX: <Unknown>
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1395 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Cucurbita maxima
: IMMEDIATE SOURCE:
: CLONE: Clone PB11
: SEQUENCE DESCRIPTION: SEQ ID NO: 1:
: US-09-734-719-1
:
: Query Match 5.5%; Score 44.6; DB 4; Length 1395;
: Best Local Similarity 63.6%; Pred.No.0.0091;
: Matches 68; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
:
: QY 692 CAGGAGGGGAAATGCTGCATTCGCTTTATATATATATATGCTTTTACGTAATTAACCTA 751
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Db 1289 CTGGAAGTTGCATTAATAATTAATTAACCTTTAATATATACATGCTTCTTACCTGCTCAAAAA 1348
:
: QY 752 AAAATATATGAACCAAAAGTAAAAAAAAAAAAAAAAAGAGAGAGAGA 798
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Db 1349 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1395
:
: RESULT 5
: US-08-232-463-14/c
: Sequence 14, Application US/08232463
: Patent No. 5670367
: GENERAL INFORMATION:

```

?	APPLICANT:	DORNER, F.
?	APPLICANT:	SCHIEFLINGER, F.
?	APPLICANT:	FALKNER, F. G.
?	TITLE OF INVENTION:	RECOMBINANT FOWLPOX VIRUS
?	NUMBER OF SEQUENCES:	52
?	CORRESPONDENCE ADDRESS:	
?	ADDRESSEE:	Foley & Lardner
?	STREET:	1800 Diagonal Road, Suite 500
?	CITY:	Alexandria
?	STATE:	VA
?	COUNTRY:	USA
?	ZIP:	2213-0299
?	COMPUTER READABLE FORM:	
?	MEDIUM TYPE:	Floppy disk
?	COMPUTER:	IBM PC compatible
?	OPERATING SYSTEM:	PC-DOS/MS-DOS
?	SOFTWARE:	PatentIn Release #1.0, Version #1.25
?	CURRENT APPLICATION DATA:	
?	APPLICATION NUMBER:	US/08/232,463
?	FILING DATE:	
?	CLASSIFICATION:	435
?	PRIOR APPLICATION DATA:	
?	APPLICATION NUMBER:	US/07/935,313
?	FILING DATE:	
?	APPLICATION NUMBER:	EP 91 114 300.6
?	FILING DATE:	26-AUG-1991
?	ATTORNEY/AGENT INFORMATION:	
?	NAME:	BENT, Stephen A.
?	REGISTRATION NUMBER:	29,768
?	REFERENCE/DOCKET NUMBER:	30472/114 IMMU
?	TELECOMMUNICATION INFORMATION:	
?	TELEPHONE:	(703)836-9300
?	TELEFAX:	(703)683-4109
?	TELEX:	899149
?	INFORMATION FOR SEQ ID NO:	14:
?	SEQUENCE CHARACTERISTICS:	
?	LENGTH:	7218 base pairs
?	TYPE:	nucleic acid
?	STRANDEDNESS:	single
?	TOPOLOGY:	linear
?	IMMEDIATE SOURCE:	
?	CLONE:	PTZgpt-F1s
?	US-08-232-463-14	
Query Match	5.5%;	Score 44.6;
Best Local Similarity	1.3%;	Pred. No. 0.019;
Matches	5;	Conservative 218; Mismatches 152; Indels 0; Gaps 0
OY	429	TGGGCGCGGAATGAGAAATCACTGACGTGACACTTACAAGCGCATCCGAATGCAGCTCAGATA 488
		::::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db	1440	TGGTCACRRR 1381
OY	469	TTAGGGGCGACCCTCAAGTTGGCTGGAGAGCATGTGGAGGAATVTCAGSGATGAACCTA 548
	::::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	1380	RRR 1321
OY	549	TGTAAATCTTTTCCCGAGAGGCTGACAAATGTAAGAACAATTTTGACATAAGCAAGAC 608
	::::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	1320	RRR 1261
OY	609	ATCTTTGTGACCATGCCCTGCACATATGCATGATGAGCTATGAACCACTGGAGAGACCCC 668
	::::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	1260	RRR 1201
OY	669	ACACGCGCTGATGATGATCACCCCCAGAGGGGAAAAATGTTGGCAATGCCTTTATATATT 728
	::::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	1200	RRR 1141
OY	729	ATGTTTACGTAATTAATGAAAAAATATGAACAACAAGTAAAAAAAAAAAAAAAAAAAAAAA 788
	::::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	1140	RRR 1081
OY	789	GAGAGAGAGAGAA 803



```

: TELEPHONE: (415) 494-8700
: TELEFAX: (415) 494-8771
: TELEX: 910 277299
: INFORMATION FOR SEQ ID NO: 17:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3214 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 220..2802
: US-08-484-105-17

Query Match 5.2%; Score 42.2; DB 1; Length 3214;
Best Local Similarity 65.3%; Pred. NO. 0.06;
Matches 62; Conservative 0; Mismatches 33; Indels 0; Gaps 0

QY 704 ATGCTGCCAATGCCCTTTATATATATGTGTTTACTGAAATTAAGTCAAAAAATATGAAA 763
Db 3119 ATTGTGAATTTTCATTAATATTAATGCTGTTTAAAAAAGAGAGAGAGAGAGAGAGAG 3178
QY 764 CCAAAAGTAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 798
Db 3179 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3213

RESULT 10
US-08-484-106-17
: Sequence 17, Application US/08484106
: Patent No. 5614618
: GENERAL INFORMATION:
: APPLICANT: STILLMAN, Bruce
: APPLICANT: BELL, Stephen P
: APPLICANT: KOBAYASHI, Ryuji
: APPLICANT: RINE, Jasper
: APPLICANT: FOSS, Margit
: APPLICANT: MCNALLY, Francis J
: APPLICANT: LAURENSEN, Patricia
: APPLICANT: HERSKOWITZ, Ira
: APPLICANT: LI, Joachim J
: APPLICANT: GAVIN, Kimberly
: TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT
: STREET: 4 Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,106
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman Ph.D., Richard Aron
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 494-8771
: TELEFAX: (415) 494-8771
: TELEX: 910 277299
: INFORMATION FOR SEQ ID NO: 17:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3214 base pairs
: TYPE: nucleic acid

```

STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 220..2802  
US-08-484-106-17

Query Match  
Best Local Similarity 5.2%; Score 42.2; DB 1; Length 3214;  
Pred. No. 0.06; Mismatches 33; Indels 0; Gaps 0;

Matches 62; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
QY 704 ATGGTGCATGCTTTATATATATGTTTCTGAAATTAATGAAATATGAAA 763  
Db 3119 ATTTGGAATTTTCATTAATTAATTTGCTTTTAAAAA 3178  
QY 764 CCAAAAGTAAAAAAGAGAGAGAGA 798  
Db 3179 AAAAAAAAAAAAAAAAAAAAAAAAAA 3213

## RESULT 11

US-09-336-643A-82  
Sequence 82, Application US/0933643A  
Patent No. 6399761  
GENERAL INFORMATION:  
APPLICANT: Miller, Andrew P.  
APPLICANT: Curran, Mark Edward  
APPLICANT: Hu, Ping  
APPLICANT: Rutger, Marc  
APPLICANT: Wang, Jian-Wang  
TITLE OF INVENTION: No. 6399761e1 Human Potassium Channels  
FILE REFERENCE: SEQ-15P  
CURRENT APPLICATION NUMBER: US/09/336,643A  
PRIOR FILING DATE: 1999-06-18  
PRIOR APPLICATION NUMBER: 60/076,687  
PRIOR FILING DATE: 1998-08-07  
PRIOR APPLICATION NUMBER: 60/116,448  
PRIOR FILING DATE: 1999-01-19  
PRIOR APPLICATION NUMBER: PCT/US99/03826  
NUMBER OF SEQ ID NOS: 87  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 82  
LENGTH: 3300  
TYPE: DNA  
ORGANISM: H. sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (50)...(1285)  
US-09-336-643A-82

Query Match  
Best Local Similarity 5.2%; Score 42.2; DB 4; Length 3300;  
Pred. No. 0.061; Mismatches 33; Indels 0; Gaps 0;

Matches 62; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
QY 708 TGGCAATGCTTTATATATATGTTTCTGAAATTAATGAAATATGAAACCA 767  
Db 3206 TGTGAATGTTGACAAATTAATGTTTAAAGTCAAAAAAAAAAAAAAAAAA 3265  
QY 768 AAGTAAAAAAGAGAGAGAGAGA 802  
Db 3266 AAAAAAAAAAAAAAAAAAAAAAAAAA 3300

## RESULT 12

US-09-385-982-389/C  
Sequence 389, Application US/09385982  
Patent No. 6262334  
GENERAL INFORMATION:  
APPLICANT: ENDEGE, WILSON O., ET AL.  
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION  
TITLE OF INVENTION: PRODUCTS: 11

FILE REFERENCE: CCDNA-260XX  
CURRENT APPLICATION NUMBER: US/09/385,982  
CURRENT FILING DATE: 1999-08-30  
EARLIER APPLICATION NUMBER: 09/328,111  
EARLIER FILING DATE: 1999-06-08  
EARLIER APPLICATION NUMBER: 60/117,393  
EARLIER FILING DATE: 1999-01-27  
EARLIER APPLICATION NUMBER: 60/098,639  
EARLIER FILING DATE: 1998-08-31  
NUMBER OF SEQ ID NOS: 544  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 389  
LENGTH: 629  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(629)  
OTHER INFORMATION: n = A,T,C or G  
US-09-385-982-389

Query Match  
Best Local Similarity 5.2%; Score 41.6; DB 4; Length 629;  
Pred. No. 0.042; Mismatches 47; Indels 0; Gaps 0;

Matches 68; Conservative 0; Mismatches 47; Indels 0; Gaps 0;  
QY 689 CCCCAGAGGGGAAATGTCGCAATGCTTTATATATATGTTTCTGAAATTAAC 748  
Db 119 CACCATGATGTGAATGTTTCAAAATTTCAAAATTAATGTTTCAAACTG 60  
QY 749 TGAATAATATGAACCAAGTAAAAAAGAGAGAGAGAGAGAA 803  
Db 59 CAAAAAAAAAAAAAAAAAAAAAAAAA 5

## RESULT 13

US-09-363-708-3  
Sequence 3, Application US/09363708  
Patent No. 6399747  
GENERAL INFORMATION:  
APPLICANT: Schmidt, et al.  
TITLE OF INVENTION: NOVEL SHC BINDING PROTEIN.  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive/6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/363,708  
FILING DATE:

CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 01017/34451  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2246 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
DESCRIPTION: /desc = "mouse PAL CDNA"





GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 20:56:58 ; Search time 74.3669 Seconds

(without alignments)  
9222.973 Million cell updates/sec

Title: US-10-082-502-16

Perfect score: 806

Sequence: 1 cggcccaagctgggagccca.....aagagagagagagaacta 806

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications\_NA:\*

- 1: /cgn2\_6/p/odata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/p/odata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/p/odata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/p/odata/1/pubpna/US07\_PUBCOMB.seq:\*
- 5: /cgn2\_6/p/odata/1/pubpna/US07\_PUBCOMB.seq:\*
- 6: /cgn2\_6/p/odata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/p/odata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/p/odata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/p/odata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/p/odata/1/pubpna/US10\_PUBCOMB.seq:\*
- 11: /cgn2\_6/p/odata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/p/odata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/p/odata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/p/odata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	750.4	93.1	814	10	US-09-864-864-309
2	749.4	93.0	1210	9	US-10-028-072-173
3	749.4	93.0	1210	9	US-10-121-049-173
4	749.4	93.0	1210	9	US-10-123-904-173
5	749.4	93.0	1210	9	US-10-140-470-173
6	749.4	93.0	1210	9	US-10-175-746-173
7	749.4	93.0	1210	9	US-10-176-918-173
8	749.4	93.0	1210	9	US-10-176-921-173
9	749.4	93.0	1210	9	US-10-137-865-173
10	749.4	93.0	1210	9	US-10-140-474-173
11	749.4	93.0	1210	9	US-10-142-431-173
12	749.4	93.0	1210	9	US-10-143-114-173
13	749.4	93.0	1210	9	US-10-140-002-173
14	749.4	93.0	1210	9	US-10-142-419-173
15	749.4	93.0	1210	9	US-10-143-262-173
16	749.4	93.0	1210	9	US-10-143-423-173
17	749.4	93.0	1210	9	US-10-121-050-173
18	749.4	93.0	1210	9	US-10-141-755-173
19	646.4	80.2	657	9	US-09-822-846-398

20	237.8	29.5	564	10	US-09-864-864-228	Sequence 228, App
21	147	18.2	147	10	US-09-954-456-779	Sequence 779, App
22	147	18.2	147	10	US-09-880-107-1335	Sequence 1335, App
23	140.2	17.4	281	10	US-09-783-590-1551	Sequence 1551, App
24	126	15.6	126	10	US-09-954-456-778	Sequence 778, App
25	126	15.6	126	10	US-09-880-107-1334	Sequence 1334, App
26	48	6.0	1864	9	US-09-809-391-130	Sequence 130, App
27	45.4	5.6	1405	10	US-09-820-003A-1	Sequence 105, App
28	45.2	5.6	1260	10	US-09-923-737-105	Sequence 128, App
29	44.8	5.6	317	10	US-10-091-483-128	Sequence 12, App
30	44.8	5.6	345	9	US-10-079-900-12	Sequence 12, App
31	44.8	5.6	545	10	US-09-764-873-112	Sequence 12, App
32	44.6	5.5	610	10	US-09-919-344-15	Sequence 14, App
33	44.6	5.5	404	10	US-09-960-352-4144	Sequence 96, App
34	44.4	5.5	520	9	US-10-091-483-96	Sequence 96, App
35	44.4	5.5	802	10	US-09-764-864-381	Sequence 381, App
36	44.4	5.5	473	10	US-09-960-352-6407	Sequence 6407, App
37	44	5.4	1837	9	US-10-227-884-39	Sequence 39, App
38	43.6	5.4	1837	9	US-10-230-163-39	Sequence 39, App
39	43.6	5.4	1837	9	US-10-218-631-39	Sequence 39, App
40	43.6	5.4	1837	9	US-10-230-338-39	Sequence 39, App
41	43.6	5.4	1837	9	US-10-230-414-39	Sequence 39, App
42	43.6	5.4	1049	10	US-09-800-729-67	Sequence 67, App
43	43.6	5.4	1834	10	US-09-934-289A-41	Sequence 41, App
44	43.4	5.4				
45	43.4	5.4				

## ALIGNMENTS

RESULT 1

US-09-864-864-309

Sequence 309, Application US/09864864

Patent No. US20020102679A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan L.

APPLICANT: Dillon, David C.

APPLICANT: Secrist, Heather

APPLICANT: Lodge, Michael J.

APPLICANT: Algate, Paul A.

APPLICANT: Fling, Steve P.

APPLICANT: Mannion, Jane

APPLICANT: Benson, Darin R.

APPLICANT: Carter, Darlick

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER

FILE REFERENCE: 210121.523

CURRENT APPLICATION NUMBER: US/09/864, 864

CURRENT FILING DATE: 2001-05-23

NUMBER OF SEQ ID NOS: 341

SOFTWARE: Corlita Invention Disclosure Database

SEQ ID NO 309

LENGTH: 814

TYPE: DNA

ORGANISM: Homo sapiens

US-09-864-864-309

Query Match

Best Local Similarity 93.1%; Score 750.4; DB 10; Length 814;

Mismatches 754; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 11 CTGGGGCAAGTGAAGTCAAGCGGCTGCGACGCGCTGGGCGGCGGCGGCGGCTGG 70

DB 52 CTGGGGCAAGTGAAGTCAAGCGGCTGCGACGCGCTGGGCGGCGGCGGCGGCTGG 111

QY 71 GACCAAGGTGAGACCAAGCGGCTTACCTTAARATGAAGGCTGGGCTGGGCTGGCCTG 130

DB 112 GAGCAGAGGTGAGGCGGCGGCGGCGGCTTACGCTTAAGATGAAGGCTGGGCTGGGCGGCTG 171

QY 131 CTTCGGGGGCGGCTGCTGGGAGACCGGCGGCTGGGAGGAGGAGGAGGAGATCTCCACTGTGGA 190

```

DB 172 CTTCTGCGGGCCCTGCTGGGAACCCGCTGGCTCGAGAGAGCCAGCATCTCCACTGTGA 231
OY 191 GCATCGAGGGCTCTGTGTGATGAAGTGAATGGAAATTTGCCAGGTGAGCCCAAGAG 250
DB 232 GCATCGAGGGCTCTGTGTGATGAAGTGAATGGAAATTTGCCAGGTGAGCCCAAGAG 291
OY 251 ACCATTGAGTGGGANTCTTCCGGATTCATCCAGATGGAGCCAGTCAGTGTGGAGTG 310
DB 292 ACCATTGAGTGGGANTCTTCCGGATTCATCCAGATGGAGCCAGTCAGTGTGGAGTG 351
OY 311 CCTTATGCCCCGCTCAGAGAGCCCACTCAGAGCTGCTGAGAGATATGTACCCGATG 370
DB 352 CCTTATGCCCCGCTCAGAGAGCCCACTCAGAGCTGCTGAGAGATATGTACCCGATG 411
OY 371 AAGAGATGSGGAAACAGATTGATCTTCCACCATTCCGAGAGTACGTACGTGTAGTG 430
DB 412 AAGAGATGSGGAAACAGATTGATCTTCCACCATTCCGAGAGTACGTACGTGTAGTG 471
OY 431 GGGCGGANTGAGGANTCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 490
DB 472 GGGCGGANTGAGGANTCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 531
OY 491 AGCGGACCCCTCAAGTTGCGTGTGAGAGCATTTGAGAGATTAAGAGATGAAGTCAAT 550
DB 532 AGCGGACCCCTCAAGTTGCGTGTGAGAGCATTTGAGAGATTAAGAGATGAAGTCAAT 591
OY 551 GAATTCCTTTCCGAGAGGCTGACATGTTAAAGACAACTTTGAGTGAAGGAAAGAT 610
DB 592 GAATTCCTTTCCGAGAGGCTGACATGTTAAAGACAACTTTGAGTGAAGGAAAGAT 651
OY 611 CTTTGTGACATGCGCTGACATATGCGATGAGCTATGAGACCACTGAGAGCCAGC 670
DB 652 CTTTGTGACATGCGCTGACATATGCGATGAGCTATGAGACCACTGAGAGCCAGC 711
OY 671 ACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730
DB 712 ACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771
OY 731 GTTTTTACTGAAATTAAGTGAAGGAAATTAAGGAAATTAAGGAAATTAAGGAAAT 772
DB 772 GTTTTTACTGAAATTAAGTGAAGGAAATTAAGGAAATTAAGGAAATTAAGGAAAT 813

```

```

PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11

```

PRIOR APPLICATION NUMBER: 60/069334  
 PRIOR FILING DATE: 1997-12-11  
 PRIOR APPLICATION NUMBER: 60/069694  
 PRIOR FILING DATE: 1997-12-16  
 PRIOR APPLICATION NUMBER: 60/072320  
 PRIOR FILING DATE: 1998-01-23  
 PRIOR APPLICATION NUMBER: 60/073612  
 PRIOR FILING DATE: 1998-02-04  
 PRIOR APPLICATION NUMBER: 60/074086  
 PRIOR FILING DATE: 1998-02-09  
 PRIOR APPLICATION NUMBER: 60/074092  
 PRIOR FILING DATE: 1998-02-09  
 PRIOR APPLICATION NUMBER: 60/077791  
 PRIOR FILING DATE: 1998-03-12  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/079294  
 PRIOR FILING DATE: 1998-03-25  
 PRIOR APPLICATION NUMBER: 60/079663  
 PRIOR FILING DATE: 1998-02-27  
 PRIOR APPLICATION NUMBER: 60/079728  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/080165  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081695  
 PRIOR FILING DATE: 1998-04-14  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081818  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082999  
 PRIOR FILING DATE: 1998-04-24  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084627  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084637  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/085149  
 PRIOR FILING DATE: 1998-05-12  
 PRIOR APPLICATION NUMBER: 60/085323  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085338  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085339  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085579  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085697  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085704  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/086414  
 PRIOR FILING DATE: 1998-05-22  
 PRIOR APPLICATION NUMBER: 60/086430  
 PRIOR FILING DATE: 1998-05-22  
 PRIOR APPLICATION NUMBER: 60/087106  
 PRIOR FILING DATE: 1998-05-28  
 PRIOR APPLICATION NUMBER: 60/088026  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088730  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088741  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088810

PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088858  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090445  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090538  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090863  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/091360  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091519  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091982  
 PRIOR FILING DATE: 1998-07-07

Query Match 93.0%; Score 749.4; DB 9; Length 1210;  
 Best Local Similarity 98.9%; Pred. No. 5.6e-200;  
 Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

11 CTGGGCGCAAGTGAAGTCCAGCGGTCTGCCAGCGCTTGGCCAGCGGGGGCGCTCG 70  
 450 CTGGGCGCAAGTGAAGTCCAGCGGTCTGCCAGCGCTTGGCCAGCGGGGGCGCTCG 509  
 71 GACCAAGGTGAGCAACCCGTTACCTAAATGAAGAGCTGGAGTGGCTGCGCTG 130  
 510 GACCAAGGTGAGCAACCCGTTACCTAAATGAAGAGCTGGAGTGGCTGCGCTG 569  
 131 CTTCCTGGGGCCCTGCTGGGAACCGCTGGGCTCGGAGAGCCAGATCTCACATGGA 190  
 570 CTTCCTGGGGCCCTGCTGGGAACCGCTGGGCTCGGAGAGCCAGATCTCACATGGA 629  
 191 GCATGAGGGCTCTGTGTGATGAATGAAGGAAATTCGCCAGGTGAGCCCAAGAG 250  
 630 GCATGAGGGCTCTGTGTGATGAATGAAGGAAATTCGCCAGGTGAGCCCAAGAG 689  
 251 ACCATTGAGTGGATCTTCCGATCAATCCAGATGAGCAGCAATGATGATGAGTGTG 310  
 690 ACCATTGAGTGGATCTTCCGATCAATCCAGATGAGCAGCAATGATGATGAGTGTG 749  
 311 CTTATGCGCGCTGAGAGCCGCTTCAAGAGCTGCTGAGAGATATGACCGGATG 370  
 750 CTTATGCGCGCTGAGAGCCGCTTCAAGAGCTGCTGAGAGATATGACCGGATG 809  
 371 AAGGATGAGTGGGAACAGATGATCTTCCAGCCATGCAAGAACTGACGTGAGTG 430  
 810 AAGGATGAGTGGGAACAGATGATCTTCCAGCCATGCAAGAACTGACGTGAGTG 869  
 431 GCGCGAATGAGAAATCAGTGAAGTGAAGCTTCAAGAGCAATGCAATGCAATGAT 490  
 870 GCGCGAATGAGAAATCAGTGAAGTGAAGCTTCAAGAGCAATGCAATGCAATGAT 929  
 491 AGCGGACCCCTCAAGTGTGCTGAGAGATGTGAGGAATATGAGAGATGAACTCAT 550  
 930 AGCGGACCCCTCAAGTGTGCTGAGAGATGTGAGGAATATGAGAGATGAACTCAT 989  
 551 GAATCTTTTCCGAGAGGCTGACAAATGTTAAAGACAACTTGCAGTAAAGGACAGAT 610  
 990 GAATCTTTTCCGAGAGGCTGACAAATGTTAAAGACAACTTGCAGTAAAGGACAGAT 1049  
 611 CTTGTGACCATGCGCTGACATATGATGATGATGATGATGATGATGATGATGATGAT 670



```
QY 71 GACCAAGCTTGGAGCAACCCGTTACCTTAARATGAAGGCTGGGTTGGCTGGCCCTG 130
DB 510 GACAGAGAGTGGAGCGACCCCATTAAGCTAAAGATGAAGGCTGGGTTGGCTGGCCCTG 509
QY 131 CTTCTGGGGGCGCTGGCTGGGGAACCGGCTGGGCTCGAGAGAGCCAGATCTCCACTGTGA 190
DB 570 CTTCTGGGGGCGCTGGCTGGGGAACCGGCTGGGCTCGAGAGAGCCAGATCTCCACTGTGA 629
QY 191 GCATGAGGAGCTCTGTGTGATGAATGAATGGAATTTGCCAGGTGGAGCCCAAGAG 250
DB 630 GCATGAGGAGCTCTGTGTGATGAATGAATGGAATTTGCCAGGTGGAGCCCAAGAG 689
QY 251 ACCATTCAGATGGGATCTTTCCGATCAATCCAGATGCGAGCCAGTCACTGGTGGAGTGTG 310
DB 690 ACCATTCAGATGGGATCTTTCCGATCAATCCAGATGCGAGCCAGTCACTGGTGGAGTGTG 749
QY 311 CCTTATGCGCGCTGAGAGGCGCCACGACAGAGCTGGTGGAGAGATATGTACCGGATG 370
DB 750 CCTTATGCGCGCTGAGAGGCGCCACGACAGAGCTGGTGGAGAGATATGTACCGGATG 809
QY 371 AAGGAGTATGGGGAACAGATTGATCTTCACCATCGGAGAACACTAGCTAGCTAGTGTG 430
DB 810 AAGGAGTATGGGGAACAGATTGATCTTCACCATCGGAGAACACTAGCTAGCTAGTGTG 869
QY 431 GCGCGGAATGGAGAACGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 490
DB 870 GCGCGGAATGGAGAACGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 929
QY 491 AGCGGCACTTCAGATTTGCGTGTGAGAGCATTTGAGAGAAATAGAGATGAAGTGAAGT 550
DB 930 AGCGGCACTTCAGATTTGCGTGTGAGAGCATTTGAGAGAAATAGAGATGAAGTGAAGT 989
QY 551 GAATCTTTTCCGAGAGGCTGACAACTTAAAGCAAACTTTGAGTGAAGCGAAGCAT 610
DB 990 GAATCTTTTCCGAGAGGCTGACAACTTAAAGCAAACTTTGAGTGAAGCGAAGCAT 1049
QY 611 CTTTGTGACATGCGCTGACATATGCGATGATGATATGAACACTGAGAGCGCCAC 670
DB 1050 CTTTGTGACATGCGCTGACATATGCGATGATGATATGAACACTGAGAGCGCCAC 1109
QY 671 ACTGGCTTGAATGAGTACCCCGAGAGGAGGAAATGCTGGCAATGCTTTTATATATAT 730
DB 1110 ACTGGCTTGAATGAGTACCCCGAGAGGAGGAAATGCTGGCAATGCTTTTATATATAT 1169
QY 731 GTTTTACTGAATTAAGTGAATTAATATGAAACCAAAAGT 771
DB 1170 GTTTTACTGAATTAAGTGAATTAATATGAAACCAAAAGT 1210
```

## RESULT 5

```
US-10-140-470-173
; Sequence 173, Application US/10140470
; Publication No. US2003002231A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
```

```
;; TITLE OF INVENTION: ACIDS ENCODING THE SAME
;; FILE REFERENCE: P3330R1C160
;; CURRENT APPLICATION NUMBER: US/10/140,470
;; PRIORITY FILING DATE: 2002-05-06
;; PRIOR APPLICATION removed - See Palm or File Wrapper
;; NUMBER OF SEQ ID NOS: 550
;; SEQ ID NO: 173
;; LENGTH: 1210
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-140-470-173
```

```
Query Match          93.0%; Score 749.4; DB 9; Length 1210;
Best Local Similarity 98.9%; Pred. No. 5,6e-200;
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 11 CTGGGGCCAAAGTGAATCCAGGCGGTGGACGGCTTGGGCGCCACGCGGCGCTGG 70
DB 450 CTGGGGCCAAAGTGAATCCAGGCGGTGGACGGCTTGGGCGCCACGCGGCGCTGG 509
QY 71 GACCAAGCTTGGAGCAACCCGTTACCTTAARATGAAGGCTGGGTTGGCTGGCCCTG 130
DB 510 GACAGAGAGTGGAGCGACCCCATTAAGCTAAAGATGAAGGCTGGGTTGGCTGGCCCTG 509
QY 131 CTTCTGGGGGCGCTGGCTGGGGAACCGGCTGGGCTCGAGAGAGCCAGATCTCCACTGTGA 190
DB 570 CTTCTGGGGGCGCTGGCTGGGGAACCGGCTGGGCTCGAGAGAGCCAGATCTCCACTGTGA 629
QY 191 GCATGAGGAGCTCTGTGTGATGAATGGAATTTGCCAGGTGGAGCCCAAGAG 250
DB 630 GCATGAGGAGCTCTGTGTGATGAATGGAATTTGCCAGGTGGAGCCCAAGAG 689
QY 251 ACCATTCAGATGGGATCTTTCCGATCAATCCAGATGCGAGCCAGTCACTGGTGGAGTGTG 310
DB 690 ACCATTCAGATGGGATCTTTCCGATCAATCCAGATGCGAGCCAGTCACTGGTGGAGTGTG 749
QY 311 CTTTATGCGCGCTGAGAGGCGCCACGACAGAGCTGGTGGAGAGATATGTACCGGATG 370
DB 750 CTTTATGCGCGCTGAGAGGCGCCACGACAGAGCTGGTGGAGAGATATGTACCGGATG 809
QY 371 AAGGAGTATGGGGAACAGATTGATCTTCACCATCGGAGAACACTAGCTAGCTAGTGTG 430
DB 810 AAGGAGTATGGGGAACAGATTGATCTTCACCATCGGAGAACACTAGCTAGCTAGTGTG 869
QY 431 GCGCGGAATGGAGAACGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 490
DB 870 GCGCGGAATGGAGAACGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 929
QY 491 AGCGGCACTTCAGATTTGCGTGTGAGAGCATTTGAGAGAAATAGAGATGAAGTGAAGT 550
DB 930 AGCGGCACTTCAGATTTGCGTGTGAGAGCATTTGAGAGAAATAGAGATGAAGTGAAGT 989
QY 551 GAATCTTTTCCGAGAGGCTGACAACTTAAAGCAAACTTTGAGTGAAGCGAAGCAT 610
DB 990 GAATCTTTTCCGAGAGGCTGACAACTTAAAGCAAACTTTGAGTGAAGCGAAGCAT 1049
QY 611 CTTTGTGACATGCGCTGACATATGCGATGATGATATGAACACTGAGAGCGCCAC 670
DB 1050 CTTTGTGACATGCGCTGACATATGCGATGATGATATGAACACTGAGAGCGCCAC 1109
QY 671 ACTGGCTTGAATGAGTACCCCGAGAGGAGGAAATGCTGGCAATGCTTTTATATATAT 730
DB 1110 ACTGGCTTGAATGAGTACCCCGAGAGGAGGAAATGCTGGCAATGCTTTTATATATAT 1169
QY 731 GTTTTACTGAATTAAGTGAATTAATATGAAACCAAAAGT 771
DB 1170 GTTTTACTGAATTAAGTGAATTAATATGAAACCAAAAGT 1210
```

## RESULT 6

```
US-10-175-746-173
; Sequence 173, Application US/10175746
; Publication No. US20030027270A1
```

```

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C353
CURRENT APPLICATION NUMBER: US/10/175,746
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 173
LENGTH: 1210
TYPE: DNA
ORGANISM: Homo Sapien
US-10-175-746-173

```

```

Query Match          93.0%; Score 749.4; DB 9; Length 1210;
Best Local Similarity 98.9%; Pred. No. 5,6e-200;
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

```

11 CTGGGGCCAAAGTGAAGTCCAGGCTGCGCAGGCTTGGGCGACGGGGGGCCCTGG 70
450 CTGGGGCCAAAGTGAAGTCCAGGCTTGGGCGACGGGGGGGGGGGGGGGGGGGG 509
71 GACCAAGGTGGAGCAACCCGTTACCTTAARATGAAGGCTGGGGTGGGCTGGCCCTG 130
510 GAGCAGAGGTGGAGGAGCCCAATTAAGTGAAGGCTGGGGTGGGCTGGGGCCG 569
131 CTCTGGGGGGGGCTGCTGGGAAACCGGCTGGGCTGGAGAGCCAGATTCACCTGTGA 190
570 CTCTGGGGGGGGCTGCTGGGAAACCGGCTGGGCTGGAGAGCCAGATTCACCTGTGA 629
191 GCATGAGGGCTCTGGTGGATGATGAGAAATGGGAAATGGCCAGGTGGACCCCAAGAG 250
630 GCATGAGGGCTCTGGTGGATGATGAGAAATGGGAAATGGCCAGGTGGACCCCAAGAG 689
251 ACCATTGAGATGGATCTTTCCGATCAATCCAGATGGCAGCCAGTGGTGGAGTGG 310
690 ACCATTGAGATGGATCTTTCCGATCAATCCAGATGGCAGCCAGTGGTGGAGTGG 749
311 CCTTATGGCCGCTCAGAGGCCCACTCAGAGAGCTGCTGGAGAGATATGTACCGGATG 370
750 CCTTATGGCCGCTCAGAGGCCCACTCAGAGAGCTGCTGGAGAGATATGTACCGGATG 809
371 AAGGATATGGGGAACAGATGATCCTTCACCATCGCAAGAGTCCGATGAGTGGAGT 430
810 AAGGATATGGGGAACAGATGATCCTTCACCATCGCAAGAGTCCGATGAGTGGAGT 869
431 GGCCGGAATGAGAAATCCAGTGAAGTCAACAGAGTCCGATGAGTGGAGTGGAGT 490
870 GGCCGGAATGAGAAATCCAGTGAAGTCAACAGAGTCCGATGAGTGGAGTGGAGT 929
491 AGCGGACACCTTAAGTGGCTGTGAGAGCATTTGGAGAGAAATACGAGATGAATCACT 550
930 AGCGGACACCTTAAGTGGCTGTGAGAGCATTTGGAGAGAAATACGAGATGAATCACT 989
551 GAATCTTTCCGAGAGGCTGACAAATGTTAAGCAAACTTTGAGTGAAGGAGCAAGT 610

```

```

DB 990 GAATCTTTCCGAGAGGCTGACAAATGTTAAGCAAACTTTGAGTGAAGGAGCAAGT 1049
QY 611 CTCTGTGACCATGCGCTGACATATGACATGATGACATGATGACATGACATGACATG 670
DB 1050 CTCTGTGACCATGCGCTGACATATGACATGATGACATGATGACATGACATGACATG 1109
QY 671 ACTGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 730
DB 1110 ACTGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1169
QY 731 GTTTTACTGAAATTAAGTGAAGAAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 771
DB 1170 GTTTTACTGAAATTAAGTGAAGAAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1210

```

```

RESULT 7
US-10-176-918-173
Sequence 173, Application US/10176918
Publication No. US20030027275A1

```

```

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin

```

```

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C382
CURRENT APPLICATION NUMBER: US/10/176,918
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 173
LENGTH: 1210
TYPE: DNA
ORGANISM: Homo Sapien
US-10-176-918-173

```

```

Query Match          93.0%; Score 749.4; DB 9; Length 1210;
Best Local Similarity 98.9%; Pred. No. 5,6e-200;
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

```

11 CTGGGGCCAAAGTGAAGTCCAGGCTGCGCAGGCTTGGGCGACGGGGGGGGGGGGGGGGGGGG 70
450 CTGGGGCCAAAGTGAAGTCCAGGCTTGGGCGACGGGGGGGGGGGGGGGGGGGG 509
71 GACCAAGGTGGAGCAACCCGTTACCTTAARATGAAGGCTGGGGTGGGCTGGCCCTG 130
510 GAGCAGAGGTGGAGGAGCCCAATTAAGTGAAGGCTGGGGTGGGCTGGGGCCG 569
131 CTCTGGGGGGGGCTGCTGGGAAACCGGCTGGGCTGGAGAGCCAGATTCACCTGTGA 190
570 CTCTGGGGGGGGCTGCTGGGAAACCGGCTGGGCTGGAGAGCCAGATTCACCTGTGA 629
191 GCATGAGGGCTCTGGTGGATGATGAGAAATGGGAAATGGCCAGGTGGAGTGGAGTGGAGT 250
630 GCATGAGGGCTCTGGTGGATGATGAGAAATGGGAAATGGCCAGGTGGAGTGGAGTGGAGT 689
251 ACCATTGAGATGGATCTTTCCGATCAATCCAGATGGCAGCCAGTGGTGGAGTGGAGT 310
690 ACCATTGAGATGGATCTTTCCGATCAATCCAGATGGCAGCCAGTGGTGGAGTGGAGT 749

```

Query Match	93.0%;	Score 749.4;	DB 9;	Length 1210;
Best Local Similarity	98.9%;	Pred. No. 5.6e-200;		
Matches 753; Conservative	1;	Mismatches 7;	Indels 0;	Gaps 0;

APPLICANT: Gerritson, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tunas, Daniel  
APPLICANT: Watanabe, Colin K.

APPLICANT: Wood,William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P330R1C154  
; CURRENT APPLICATION NUMBER: US/10/137,865  
; CURRENT FILING DATE: 2002-05-03  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-137-865-173

Query Match 93.0%; Score 749.4; DB 9; Length 1210;  
Best Local Similarity 98.9%; Pred. No. 5.6e-200;  
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

11 CTGGGGCCAAAGTGAATCCAGCGCTGCGCAGCGCTTGGGCCACGGCGGCGCCCTGG 70  
450 CTGGGGCCAAAGTGAATCCAGCGCTGCGCAGCGCTTGGGCCACGGCGGCGCCCTGG 509  
71 GACCAAGGTGAGACACCCGCTACCTAAABATGAAAGGCTGGGGTTGGCTGGCCCTG 130  
510 GAGCAGAGGTGAGACACCCGCTACCTAAABATGAAAGGCTGGGGTTGGCTGGCCCTG 569  
131 CTCTGGGGGGCCCTGCTGGGAAACCCGCTGGGCTCGGAGAGCCAGATCTCCACTGTGA 190  
570 CTCTGGGGGGCCCTGCTGGGAAACCCGCTGGGCTCGGAGAGCCAGATCTCCACTGTGA 629  
191 GCATCAGAGGCTCTGTGGATGAAATAGAAATGGGAAATTTGCCAGGTGAGCCCAAGAAG 250  
630 GCATCAGAGGCTCTGTGGATGAAATAGAAATGGGAAATTTGCCAGGTGAGCCCAAGAAG 689  
251 ACCATTGAGTGGATCTTTCGCGATCAATCCAGATGGAGCCAGTCACTGAGTGGAGTG 310  
690 ACCATTGAGTGGATCTTTCGCGATCAATCCAGATGGAGCCAGTCACTGAGTGGAGTG 749  
311 CCTTATGGCCGCTCAGAGGCCCACTCAAGAGCTGTGAGAGAGATATGTGACCGGATG 370  
750 CCTTATGGCCGCTCAGAGGCCCACTCAAGAGCTGTGAGAGAGATATGTGACCGGATG 809  
371 AAGGATATGGGGAACAGATTGATCTTCCACCCATGCCAAGAACTAGTACGTAGTG 430  
810 AAGGATATGGGGAACAGATTGATCTTCCACCCATGCCAAGAACTAGTACGTAGTG 869  
431 GGCAGGAATGGAATCCAGTGAACCTGAGACCTACAAAGCATCCGATCGACTCAGATATT 490  
870 GGCAGGAATGGAATCCAGTGAACCTGAGACCTACAAAGCATCCGATCGACTCAGATATT 929  
491 AGCGCACCTCTCAAGTTTGGCTGTGAGAGCATTTGTGAGAGAAATACGAGATGAATCAATT 550  
930 AGCGCACCTCTCAAGTTTGGCTGTGAGAGCATTTGTGAGAGAAATACGAGATGAATCAATT 989  
551 GAATCTTTTCCCGAGAGGCTGACAAATGTTAAAGCAAACTTTGGATGAAGGAACAAT 610  
990 GAATCTTTTCCCGAGAGGCTGACAAATGTTAAAGCAAACTTTGGATGAAGGAACAAT 1049  
611 CTCTGTGACATGCGCTCAGATATCGATGATGAGTATGAAACCATGAGACCCAC 670  
1050 CTCTGTGACATGCGCTCAGATATCGATGATGAGTATGAAACCATGAGACCCAC 1109  
671 ACTGGCTTGAATGATCACCACCCAGAGAGGGAATGTGGCAATGCCCTTTATATATTAT 730  
1110 ACTGGCTTGAATGATCACCACCCAGAGAGGGAATGTGGCAATGCCCTTTATATATTAT 1169  
731 GTTTTACTGAATTAATGAATAAATATGAACCAAAAGT 771  
1170 GTTTTACTGAATTAATGAATAAATATGAACCAAAAGT 1210

RESULT 10

US-10-140-474-173

; Sequence 173, Application US/10140474  
; Publication No. US20030032156A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Mei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gunney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Watanabe, Collin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P330R1C162  
; CURRENT APPLICATION NUMBER: US/10/140,474  
; CURRENT FILING DATE: 2002-05-06  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-140-474-173

Query Match 93.0%; Score 749.4; DB 9; Length 1210;  
Best Local Similarity 98.9%; Pred. No. 5.6e-200;  
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

11 CTGGGGCCAAAGTGAATCCAGCGCTGCGCAGCGCTTGGGCCACGGCGGCGCCCTGG 70  
450 CTGGGGCCAAAGTGAATCCAGCGCTGCGCAGCGCTTGGGCCACGGCGGCGCCCTGG 509  
71 GACCAAGGTGAGACACCCGCTACCTAAABATGAAAGGCTGGGGTTGGCTGGCCCTG 130  
510 GAGCAGAGGTGAGACACCCGCTACCTAAABATGAAAGGCTGGGGTTGGCTGGCCCTG 569  
131 CTCTGGGGGGCCCTGCTGGGAAACCCGCTGGGCTCGGAGAGCCAGATCTCCACTGTGA 190  
570 CTCTGGGGGGCCCTGCTGGGAAACCCGCTGGGCTCGGAGAGCCAGATCTCCACTGTGA 629  
191 GCATCAGAGGCTCTGTGGATGAAATAGAAATGGGAAATTTGCCAGGTGAGCCCAAGAAG 250  
630 GCATCAGAGGCTCTGTGGATGAAATAGAAATGGGAAATTTGCCAGGTGAGCCCAAGAAG 689  
251 ACCATTGAGTGGATCTTTCGCGATCAATCCAGATGGAGCCAGTCACTGAGTGGAGTG 310  
690 ACCATTGAGTGGATCTTTCGCGATCAATCCAGATGGAGCCAGTCACTGAGTGGAGTG 749  
311 CCTTATGGCCGCTCAGAGGCCCACTCAAGAGCTGTGAGAGAGATATGTGACCGGATG 370  
750 CCTTATGGCCGCTCAGAGGCCCACTCAAGAGCTGTGAGAGAGATATGTGACCGGATG 809  
431 GGCAGGAATGGAATCCAGTGAACCTGAGACCTACAAAGCATCCGATCGACTCAGATATT 490  
870 GGCAGGAATGGAATCCAGTGAACCTGAGACCTACAAAGCATCCGATCGACTCAGATATT 929  
491 AGCGCACCTCTCAAGTTTGGCTGTGAGAGCATTTGTGAGAGAAATACGAGATGAATCAATT 550  
930 AGCGCACCTCTCAAGTTTGGCTGTGAGAGCATTTGTGAGAGAAATACGAGATGAATCAATT 989



```

QY 551 GAATCTTTTCCGAGAGCTGACATGTTAAAGCAAACTTGGAGTACGACAGAT 610
| | | | |
DB 990 GAATCTTTTCCGAGAGCTGACATGTTAAAGCAAACTTGGAGTACGACAGAT 1049
| | | | |
QY 611 CTTTGTGACATGCGCTGACATATCGCATATGATGATTAACACCTGGAGCCAC 670
| | | | |
DB 1050 CTTTGTGACATGCGCTGACATATCGCATATGATGATTAACACCTGGAGCCAC 1109
| | | | |
QY 671 ACTGCGTGTGATGATCACCAGAGAGGGAATGTCGCAATGCTTTATATATAT 730
| | | | |
DB 1110 ACTGCGTGTGATGATCACCAGAGAGGGAATGTCGCAATGCTTTATATATAT 1169
| | | | |
QY 731 GTTTTCTGAAATTAATGAAATTAATGAAACCAAAAGT 771
| | | | |
DB 1170 GTTTTCTGAAATTAATGAAATTAATGAAACCAAAAGT 1210
| | | | |

```

```

RESULT 11
; Sequence 173, Application US/10142431
; Publication No. US2003036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Fliviaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumanabe, Colin K
; APPLICANT: Watanabe, Daniel
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-431-173

```

```

Query Match 93.0%; Score 749.4; DB 9; Length 1210;
Best Local Similarity 98.9%; Pred. No. 5,6e-200;
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 11 CTTGGGCGCAAGTGAAGTCCAGCGCTTGGCGAGCGCTGGCCACGGCGGCGCCCTGG 70
| | | | |
DB 450 CTTGGGCGCAAGTGAAGTCCAGCGCTTGGCGAGCGCTGGCCACGGCGGCGCCCTGG 509
| | | | |
QY 71 GACCAAAAGTGAAGCAACCCGTTACCTTAARATGAAGGCTGGGCTGCTGGCCCTG 130
| | | | |
DB 510 GACCAAGAGTGAAGCAACCCGTTACCTTAARATGAAGGCTGGGCTGCTGGCCCTG 569
| | | | |
QY 131 CTTTGGGGGCGCTGCTGGGAACCCGCTGGGCTGGAGAGCGCAGATCTCACTGTGA 190
| | | | |
DB 570 CTTTGGGGGCGCTGCTGGGAACCCGCTGGGCTGGAGAGCGCAGATCTCACTGTGA 629
| | | | |
QY 191 GCATCAGAGGCTCTGTGGATGAATAGATGGAATTTGCCAGGTGAGACCCCAAGAG 250
| | | | |
DB 630 GCATCAGAGGCTCTGTGGATGAATAGATGGAATTTGCCAGGTGAGACCCCAAGAG 689
| | | | |

```

```

QY 251 ACCATGAGATGGGATCTTTCCGATATCATTCAGATGGCAGCCAGTATGTGTGAGATG 310
| | | | |
DB 690 ACCATGAGATGGGATCTTTCCGATATCATTCAGATGGCAGCCAGTATGTGTGAGATG 749
| | | | |
QY 311 CTTATGCCCCGCTCAGAGGCCACCTCAGAGAGCTGCTGGAGGAGATATGATGACCCGATG 370
| | | | |
DB 750 CTTATGCCCCGCTCAGAGGCCACCTCAGAGAGCTGCTGGAGGAGATATGATGACCCGATG 809
| | | | |
QY 371 AAGAGATATGGGAGCAACAGATGATCTTCCACCATTCGCAAGAACTACGATGCTAGTG 430
| | | | |
DB 810 AAGAGATATGGGAGCAACAGATGATCTTCCACCATTCGCAAGAACTACGATGCTAGTG 869
| | | | |
QY 431 GCGCGAATGAGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 490
| | | | |
DB 870 GCGCGAATGAGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 929
| | | | |
QY 491 AGCGGACCTCAAGTTGGGTGTGAGAGCATTTGTGAGAGATACGAGATGCAATG 550
| | | | |
DB 930 AGCGGACCTCAAGTTGGGTGTGAGAGCATTTGTGAGAGATACGAGATGCAATG 989
| | | | |
QY 551 GAATCTTTTCCGAGAGCTGACATGTTAAAGCAAACTTGGACGTAAAGCAACAGAT 610
| | | | |
DB 990 GAATCTTTTCCGAGAGCTGACATGTTAAAGCAAACTTGGACGTAAAGCAACAGAT 1049
| | | | |
QY 611 CTTTGTGACATGCGCTGACATATCGCATATGATGATTAACACCTGGAGCCAC 670
| | | | |
DB 1050 CTTTGTGACATGCGCTGACATATCGCATATGATGATTAACACCTGGAGCCAC 1109
| | | | |
QY 671 ACTGCGTGTGATGATCACCAGAGAGGGAATGTCGCAATGCTTTATATATAT 730
| | | | |
DB 1110 ACTGCGTGTGATGATCACCAGAGAGGGAATGTCGCAATGCTTTATATATAT 1169
| | | | |
QY 731 GTTTTCTGAAATTAATGAAATTAATGAAACCAAAAGT 771
| | | | |
DB 1170 GTTTTCTGAAATTAATGAAATTAATGAAACCAAAAGT 1210
| | | | |

```

```

RESULT 12
; Sequence 173, Application US/10143114
; Publication No. US2003036180A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Fliviaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumanabe, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C211
; CURRENT APPLICATION NUMBER: US/10/143,114
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-143-114-173

```

```

Query Match 93.0%; Score 749.4; DB 9; Length 1210;

```



RESULT 14  
US-10-142-419-173

Sequence 173, Application US/10142419  
Publication No. US20030044945A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C244  
CURRENT APPLICATION NUMBER: US/10/142,419  
CURRENT FILING DATE: 2002-05-10  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-142-419-173

Query Match 93.0%; Score 749.4; DB 9; Length 1210;  
Best Local Similarity 98.9%; Pred. No. 5,6e-200;  
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

11 CTGGGGCCAAAGTGAAGTCCAGGCGTCTGCGAGCGCTTGGGCGACGGGGGCGGCGCTGG 70  
450 CTGGGGCCAAAGTGAAGTCCAGGCGTCTGCGAGCGCTTGGGCGACGGGGGCGGCGCTGG 509  
71 GACCAAAAGGTGGAGCAACCCCGTTACCTTAARATGAAGGCTGGGGTGGCTGGCCCTG 130  
510 GAGCAGAGGTGGAGCGACCCATTACGCTAAAGATGAAGGCTGGGGTGGCTGGCCCTG 569  
131 CTCTGGGGGCGCTGCGGGGAACCGGCTGGGCTCGGAGGAGCCAGATCTCCACTGTGA 190  
570 CTCTGGGGGCGCTGCGGGGAACCGGCTGGGCTCGGAGGAGCCAGATCTCCACTGTGA 629  
191 GCATGAGGGGCTGCTGGTGAATGATGAGTGGAAATTCGCCAGGAGGAGCCCAAGAG 250  
630 GCATGAGGGGCTGCTGGTGAATGATGAGTGGAAATTCGCCAGGAGGAGCCCAAGAG 689  
251 ACATTCAGATGGGATCTTCCGATCAATCCAGATGGAGCCAGTCAAGTGGTGAAGTG 310  
690 ACATTCAGATGGGATCTTCCGATCAATCCAGATGGAGCCAGTCAAGTGGTGAAGTG 749  
311 CCTATAGCCGCTCAGAGGCGCACTCAGAGCTGCTGGAGAGAGATATGTGACCGGATG 370  
750 CCTATAGCCGCTCAGAGGCGCACTCAGAGCTGCTGGAGAGAGATATGTGACCGGATG 809  
371 AAGAGATATGGGAGACAGATGATCTTCCAGCCATCGCAAGAACTACAGTACTGTAGT 430  
810 AAGAGATATGGGAGACAGATGATCTTCCAGCCATCGCAAGAACTACAGTACTGTAGT 869  
431 GGGCGGAATGGAGAAATTCAGTGAATCGAGCTTCAAGGCAATCCGAATCGACTGATAT 490  
870 GGGCGGAATGGAGAAATTCAGTGAATCGAGCTTCAAGGCAATCCGAATCGACTGATAT 929

491 AGCGGCAACCTCAAGTTTGGCTGTGAGAGCATTTGTGGAGGAATAGAGGATGAATCTATT 550  
930 AGCGGCAACCTCAAGTTTGGCTGTGAGAGCATTTGTGGAGGAATAGAGGATGAATCTATT 989  
551 GAATCTTTTCCGAGAGGCTGACATGTTAAAGACAACTTTGACAGTAAGCGAAGCAT 610  
990 GAATCTTTTCCGAGAGGCTGACATGTTAAAGACAACTTTGACAGTAAGCGAAGCAT 1049  
611 CTCTGGAGCAATGCCGCGACATATCCGANGATGAGCTATGAACCACTGGAGCGCCAC 670  
1050 CTCTGGAGCAATGCCGCGACATATCCGANGATGAGCTATGAACCACTGGAGCGCCAC 1109  
671 ACTGGCTTGTATGATCAACCCCGAGAGGGGAAATGGTGGCAATGCTTTATATATAT 730  
1110 ACTGGCTTGTATGATCAACCCCGAGAGGGGAAATGGTGGCAATGCTTTATATATAT 1169  
731 GTTTTCTGAAATTAATGAAAAATATGAACCAAAAGT 771  
1170 GTTTTCTGAAATTAATGAAAAATATGAACCAAAAGT 1210

RESULT 15  
US-10-123-262-173

Sequence 173, Application US/10123262  
Publication No. US20030049816A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C38  
CURRENT APPLICATION NUMBER: US/10/123,262  
CURRENT FILING DATE: 2002-04-15  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-123-262-173

Query Match 93.0%; Score 749.4; DB 9; Length 1210;  
Best Local Similarity 98.9%; Pred. No. 5,6e-200;  
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

11 CTGGGGCCAAAGTGAAGTCCAGGCGTCTGCGAGCGCTTGGGCGACGGGGGCGGCGCTGG 70  
450 CTGGGGCCAAAGTGAAGTCCAGGCGTCTGCGAGCGCTTGGGCGACGGGGGCGGCGCTGG 509  
71 GACCAAAAGGTGGAGCAACCCCGTTACCTTAARATGAAGGCTGGGGTGGCTGGCCCTG 130  
510 GACCAAAAGGTGGAGCAACCCCGTTACCTTAARATGAAGGCTGGGGTGGCTGGCCCTG 569  
131 CTCTGGGGGCGCTGCTGGGAAACCGGCTGGGCTCGGAGGAGCCAGAGTCTCCACTGTGA 190  
570 CTCTGGGGGCGCTGCTGGGAAACCGGCTGGGCTCGGAGGAGCCAGAGTCTCCACTGTGA 629  
191 GCATGAGGGGCTGCTGGTGAATGATGAGTGGAAATTCGCCAGGTGAGCCCAAGAG 250

```

Db 630 GATGAGGCTCTGTGGATGAACAGATGGAAATGCCCAGGTGACCCCAAGAAG 689
QY 251 ACCATTACAGATGGATCTTCCGATCAATCCAGATGGCAGCCAGTCAGTGGAGTG 310
Db 690 ACCATTACAGATGGATCTTCCGATCAATCCAGATGGCAGCCAGTCAGTGGAGTG 749
QY 311 CCTTATGCCCGCTCAGAGGCCCACTCAGAGAGCTGCTGGAGAGATATGTACCGGATG 370
Db 750 CCTTATGCCCGCTCAGAGGCCCACTCAGAGAGCTGCTGGAGAGATATGTACCGGATG 809
QY 371 AAGGAGTATGGGAAACAGATGATCCTTCCACCATCCGAAGAACATACGTACGTAGTG 430
Db 810 AAGGAGTATGGGAAACAGATGATCCTTCCACCATCCGAAGAACATACGTACGTAGTG 869
QY 431 GGCCGGAATGAGAAATCCAGTAACTGAGACTACAAGGCATCCGAATCGATCAGATATT 490
Db 870 GGCCGGAATGAGAAATCCAGTAACTGAGACTACAAGGCATCCGAATCGATCAGATATT 929
QY 491 AGCGGCACCCCTCAAGTTGCGTGTGAGAGCAATGTGAGAGAAATACGAGGATGAATCATTT 550
Db 930 AGCGGCACCCCTCAAGTTGCGTGTGAGAGCAATGTGAGAGAAATACGAGGATGAATCATTT 989
QY 551 GAATCTTTTCCGAGAGGCTGACAAATGTTAAAGACAAACTTTGCAGTAAAGCAACAGAT 610
Db 990 GAATCTTTTCCGAGAGGCTGACAAATGTTAAAGACAAACTTTGCAGTAAAGCAACAGAT 1049
QY 611 CTTTGTACCATGACCTGACATATCGCATGTGAGCTATGAACCACTGAGAGCCGAC 670
Db 1050 CTTTGTACCATGACCTGACATATCGCATGTGAGCTATGAACCACTGAGAGCCGAC 1109
QY 671 ACTGGCTTATGATGATCACCCTCAGAGGGGAAATGCTGCAATGCTTTTATATATTAT 730
Db 1110 ACTGGCTTATGATGATCACCCTCAGAGGGGAAATGCTGCAATGCTTTTATATATTAT 1169
QY 731 GTTTTACTGAATTTAACTGAATAAATATGAATAACCAAAAGT 771
Db 1170 GTTTTACTGAATTTAACTGAATAAATATGAATAACCAAAAGT 1210

```

Search completed: March 30, 2003, 22:23:35  
 Job time : 86.3669 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 18:36:03 ; Search time 1147.74 Seconds

(Without alignments)  
11373.235 Million cell updates/sec

Title: US-10-082-502-16

Sequence: 1 cggcccaagctcgggccca.....aagagagagagagagacta 806

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rnd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	731	90.7	1117	13	BM552907
2	714.2	88.6	783	13	BM550436
3	713.2	88.5	868	14	BM550436
4	705.4	87.5	880	14	BM550436
5	698.6	86.7	875	13	BM550436
6	693	86.0	773	13	BM550162

7	676.6	83.9	742	12	BM522459	BM522459
8	674.4	83.7	742	12	BM522459	BM522459
9	671.4	83.3	867	13	BM522459	BM522459
10	669.8	83.1	673	14	BM522459	BM522459
11	660.8	82.0	720	13	BM522459	BM522459
12	650.8	80.7	791	12	BM522459	BM522459
13	646.4	80.2	791	12	BM522459	BM522459
14	645	80.0	765	12	BM522459	BM522459
15	641	79.5	745	12	BM522459	BM522459
16	639	79.3	818	13	BM522459	BM522459
17	636.8	79.0	783	13	BM522459	BM522459
18	628.2	77.9	719	12	BM522459	BM522459
19	623.8	77.4	746	13	BM522459	BM522459
20	620	76.9	856	12	BM522459	BM522459
21	616.6	76.5	751	12	BM522459	BM522459
22	614.6	76.3	699	9	AL548119	AL548119
23	607.8	75.4	930	14	BM522459	BM522459
24	604.2	75.0	611	13	BM522459	BM522459
25	603.6	74.9	853	12	BM522459	BM522459
26	602.2	74.7	797	13	BM522459	BM522459
27	599	74.3	599	13	BM522459	BM522459
28	590.6	73.3	591	14	BM522459	BM522459
29	589.8	73.2	593	10	BM522459	BM522459
30	587.6	72.9	588	14	BM522459	BM522459
31	586.8	72.8	885	14	BM522459	BM522459
32	586.8	72.8	963	14	BM522459	BM522459
33	585.4	72.6	1019	12	BM522459	BM522459
34	584.4	72.5	908	14	BM522459	BM522459
35	583.8	72.4	1255	11	AK019927	AK019927
36	583.2	72.4	772	11	AK019927	AK019927
37	583	72.3	892	13	BM522459	BM522459
38	580.4	72.0	776	13	BM522459	BM522459
39	578	71.7	587	10	BE042694	BE042694
40	577.2	71.6	754	11	AK013568	AK013568
41	574	71.2	853	11	AK013014	AK013014
42	572	71.0	594	14	BM020508	BM020508
43	571.6	70.9	587	14	AI860382	AI860382
44	570.4	70.8	808	13	BI408525	BI408525
45	570	70.7	598	9	AI418940	AI418940

## ALIGNMENTS

RESULT 1  
LOCUS BM552907  
DEFINITION AGENCOURT.6572552 NIH\_MGC\_41 Homo sapiens CDNA clone IMAGE:5466942  
5' mRNA sequence.  
ACCESSION BM552907  
VERSION BM552907.1 GI:18791172  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1117)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: DCC/DTP  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution by: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LCM1968 row: C column: 07  
High quality sequence stop: 709.  
Location/Qualifiers 1. 1117

FEATURES  
SOURCE

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5466942"
/clone_1lb="NIH_MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT      299 a      258 c      320 g      239 t      1 others
ORIGIN

```

```

Query Match      90.7%; Score 731; DB 13; Length 1117;
Best Local Similarity 95.5%; Pred. No. 5e-119;
Matches 762; Conservative 1; Mismatches 32; Indels 3; Gaps 1;

```

```

Oy 11 CTGGGCGCAAGTGAAGTCCAGCGCTGCCAGCGCTGGGCGCAGCGCGCGCTGG 70
Db 55 CTGGGCGCAAGTGAAGTCCAGCGCTGCCAGCGCTGGGCGCAGCGCGCGCTGG 114
Oy 71 GACCAAGGTGGAGCAACCCGCTTACCTAAATGAAAGGCTGGGGTGGCTGCCCTG 130
Db 115 GAGCAGAGGTGGAGCGACCCCATTCAGCTAAAGTAAAGGCTGGGGTGGCTGCCCTG 174
Oy 131 CTTCGGGGGGCGCTCTGGGAAACCGCGTGGGCTGGAGAGCGCAGATCTCCACTGTGA 190
Db 175 CTTCGGGGGGCGCTCTGGGAAACCGCGTGGGAGAGCGCAGATCTCCACTGTGA 234
Oy 191 GCATGCGAGGCTCTGGTGGATGAATAGATGGAATGGAAATGGCCAGGTGAGCCCAAGAAG 250
Db 235 GCATGCGAGGCTCTGGTGGATGAATAGATGGAATGGAAATGGCCAGGTGAGCCCAAGAAG 294
Oy 251 ACCATTGAGTGGGATCTTTCCGGATCAATCAGATGGCAGCAGTCACTGGTGGAGTG 310
Db 295 ACCATTGAGTGGGATCTTTCCGGATCAATCAGATGGCAGCAGTCACTGGTGGAGTG 354
Oy 311 CCTTATGCGCGCTCAGAGCGCCACTCAGAGCTGCTGGAGAGATATGTCACCGGATG 370
Db 355 CCTTATGCGCGCTCAGAGCGCCACTCAGAGCTGCTGGAGAGATATGTCACCGGATG 414
Oy 371 AAGGATGAGGGAACAGATTGATCTTCCACCCATCCAGAACTAGCTAGCTAGTG 430
Db 415 AAGGATGAGGGAACAGATTGATCTTCCACCCATCCAGAACTAGCTAGCTAGTG 474
Oy 431 GCGCGGATGGAATCCAGTGAATGAGTGAACCTCAAGGATCCGAATGACTCAATATT 490
Db 475 GCGCGGATGGAATCCAGTGAATGAGTGAACCTCAAGGATCCGAATGACTCAATATT 534
Oy 491 AGCGCACCCCTCAAGTTCGCTGAGAGCATTTGGAGAAATAGAGAGTAATCTATT 550
Db 535 AGCGCACCCCTCAAGTTCGCTGAGAGCATTTGGAGAAATAGAGAGTAATCTATT 594
Oy 551 GAATTTCTTTCCGAGAGGCTGACAAATGTTAAAGCAAACTTTGCAATAGCAAGACAGAT 610
Db 595 GAATTTCTTTCCGAGAGGCTGACAAATGTTAAAGCAAACTTTGCAATAGCAAGACAGAT 654
Oy 611 CTTTGTGACATGCGCTGACATATCGATGATGAGTATGACACCTGGAGACGCCAC 670
Db 655 CTTTGTGACATGCGCTGACATATCGATGATGAGTATGACACCTGGAGACGCCAC 714
Oy 671 ACTGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 727
Db 715 ACTGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774
Oy 728 TATGTTTTTACTGAATTAATGAAAAATATGAAACCAAAAGTAAAAAAGTAAAAAAGT 787
Db 775 TGGTTTTTACTGAAATTAATGAAAAATATGAAACCAAAAGTAAAAAAGTAAAAAAGT 834

```

```

Oy 788 AGAGAGAGAGAGAGACT 805
Db 835 AAAAAAAGTGAAGGCAATT 852

```

```

RESULT 2
LOCUS      B1550436      783 bp      mRNA      linear      EST 05-SEP-2001
DEFINITION 603192905P1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5264328 5',
            mRNA sequence.
ACCESSION  B1550436
VERSION    B1550436.1 GI:15437748
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 783)
            NIH-MGC http://mgc.ncl.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            JOURNAL  Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgrsbs@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLM11666 row: 1 column: 01
            High quality sequence stop: 730.
            Location/Qualifiers

```

```

1. 783
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5264328"
/clone_1lb="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/Note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.5 kb and
normalized to R0F 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      209 a      189 c      225 g      160 t
ORIGIN

```

```

Query Match      88.6%; Score 714.2; DB 13; Length 783;
Best Local Similarity 97.1%; Pred. No. 5.3e-116;
Matches 747; Conservative 1; Mismatches 19; Indels 2; Gaps 2;

```

```

Oy 21 AGTGAAGTCCAGCGCTTGGCCAGCGCTTGGGCGCGCGCGCTGGAGCAAGAT 80
Db 1 AGCGGATTCACAGCGCTTTCACAGCGCTTGGGCGCGCGCGCTGGAGCAAGAT 60
Oy 81 GGAGCAACCCGTTACCTTAATGAAGAGCTGGGGTTGGCTGGCTGCTTCTGGGG 140
Db 61 GGAGCAACCCATTCATGTAAGAGAGCTGGGGTTGGCTGGCTGCTTCTGGGG 120
Oy 141 CCTGCTGGGAACGCGCTGGGCTGGAGAGCAGAGATCTCAGTGTGAGATGACAGG 200
Db 121 CCTGCTGGGAACGCGCTGGGCTGGAGAGCAGAGATCTCAGTGTGAGATGACAGG 180
Oy 201 CTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 260
Db 181 CTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240

```

```

Oy 261 TGGGATCTTCCGATCAATCCAGATGGACACAGTCACTGCTGGAGTGCCTTATGCC 320
Db 241 TGGGATCTTCCGATCAATCCAGATGGACACAGTCACTGCTGGAGTGCCTTATGCC 300
Oy 321 GCTCAGAGGCCCCACCTCAGAGAGTGTGAGAGATATGAGCCGATGAAGAGATATG 380
Db 301 GCTAAGAGGCCCCACCTCAGAGAGTGTGAGAGATATGAGCCGATGAAGAGATATG 360
Oy 381 GGGAAAGAGATGATCTCTCCACCCATCGCAAGAACTACGATGATGATGAGGCGGATG 440
Db 361 GGGAAAGAGATGATCTCTCCACCCATCGCAAGAACTACGATGATGATGAGGCGGATG 420
Oy 441 GAGAATCCAGATGATGATCTCTCCACCCATCGCAAGAACTACGATGATGATGAGGCGGATG 500
Db 421 GAGAATCCAGATGATGATCTCTCCACCCATCGCAAGAACTACGATGATGATGAGGCGGATG 480
Oy 501 TCAAGTTGGGTGTGAGAGATGTTGAGAGAACTACGATGATGATGATGAGGCGGATG 560
Db 481 TCAAGTTGGGTGTGAGAGATGTTGAGAGAACTACGATGATGATGATGAGGCGGATG 540
Oy 561 CCCGAGAGGCTGACATGTTAAAGACAACTTTGCACTAAGCGAAGATCTTTGTGACC 620
Db 541 CCCGAGAGGCTGACATGTTAAAGACAACTTTGCACTAAGCGAAGATCTTTGTGACC 600
Oy 621 ATGCCCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 680
Db 601 ATGCCCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 659
Oy 681 TGATCAGACCCCGAGAGAGGAGAAATGTTGCAATGACC-TTTTATATATATGTTTACT 739
Db 660 TGATCAGACCCCGAGAGAGGAGAAATGTTGCAATGACC-TTTTATATATATGTTTACT 719
Oy 740 GAATTTACTGAAAAATATGAAACCAAAAGTAAAAAATGAAAAAATGAAAAAATGAAAAA 788
Db 720 GAATTTACTGAAAAATATGAAACCAAAAGTAAAAAATGAAAAAATGAAAAAATGAAAAA 768

RESULT 3
LOCUS B0421069 868 bp mRNA linear EST 23-MAY-2002
DEFINITION AGENCOURT_7911956 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6009949
5', mRNA sequence.
ACCESSION B0421069
VERSION B0421069.1 GI:21116384
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 868)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaab@remail.nih.gov
Tissue Procurement: DCTD/DP/GenDAR
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1M13197 row: d column: 14
High quality sequence stop: 542.
Location/Qualifiers
1..868
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6009949"
/clone_11b="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: PCMV-SPORE6; Site_1: NotI;

```

```

BASE COUNT 232 a 203 c 257 g 168 t 8 others
ORIGIN
Query Match 88.5%; Score 713.2; DB 14; Length 868;
Best Local Similarity 94.6%; Pred. No. 7 6e-116;
Matches 755; Conservative 1; Mismatches 37; Indels 5; Gaps 2;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies.
Oy 11 CTGGGGCCAAAGTAAAGTCCAGGCTGTCAGAGCCCTTGGGCGACGCGGCGCTGG 70
Db 37 CTGGGGCCAAAGTAAAGTCCAGGCTGTCAGAGCCCTTGGGCGACGCGGCGCTGG 96
Oy 71 GACCAAGGAGGAGCAACCCGTTACCTTAAATGAAAGCTGGGTTGGTGGCCCTGG 130
Db 97 GAGCAGAGGTGGAGCGACCCCATTTACCTTAAAGTAAAGCTGGGTTGGTGGCCCTGG 156
Oy 131 CTCTGGGGGCGCTGCTGGGAAACCGCTGGGCTCGAGAGCCAGATCTCACTGGAG 190
Db 157 CTCTGGGGGCGCTGCTGGGAAACCGCTGGGCTCGAGAGCCAGATCTCACTGGAG 216
Oy 191 GCATGACGGGCTGTGTGATGAATGAATGGAATTTGCCAGGTGACCCCAAGAG 250
Db 217 GCATGACGGGCTGTGTGATGAATGAATGGAATTTGCCAGGTGACCCCAAGAG 276
Oy 251 ACCATTGAGATGAGATCTTCCGGATCAATCCAGATGCGCAGATGATGATGATGATG 310
Db 277 ACCATTGAGATGAGATCTTCCGGATCAATCCAGATGCGCAGATGATGATGATGATG 336
Oy 311 CCTTATGCGCGCTCAGAGGCGCCACTCACAGAGCTCTGGAGAGATATGACCGGATG 370
Db 337 CCTTATGCGCGCTCAGAGGCGCCACTCACAGAGCTCTGGAGAGATATGACCGGATG 396
Oy 371 AAGAGATATGAGGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATG 430
Db 397 AAGAGATATGAGGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATG 456
Oy 431 GCGCGAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 490
Db 457 GCGCGAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 516
Oy 491 AGCGGACCCCTCAAGTTTGGTGTGAGAGCATTTGGAGAAATACAGAGATGATGATG 550
Db 517 AGCGGACCCCTCAAGTTTGGTGTGAGAGCATTTGGAGAAATACAGAGATGATGATG 576
Oy 551 GAATCTTTTCCGAGAGAGGCTGCAATGTTAAAGACAACTTTGCACTAAGCGAAGAT 610
Db 577 GAATCTTTTCCGAGAGAGGCTGCAATGTTAAAGACAACTTTGCACTAAGCGAAGAT 636
Oy 611 CTCTGTGACCATGCGCTGACATATGATGATGATGATGATGATGATGATGATGATGATG 670
Db 637 CTCTGTGACCATGCGCTGACATATGATGATGATGATGATGATGATGATGATGATGATG 696
Oy 671 ACTGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 728
Db 697 ACTGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 756
Oy 729 ATG--TTTCTGTAATAATTAAGTAAATATGAAACCAAAAGTAAAAAATGAAAAA 785
Db 757 ATGTTTATTTACCGGAAATTAACCTGAAAAAATATGAAACCAAAAGTAAAAAATGAAAAA 816
Oy 786 AAAGAGAGAGAGAGAGAA 803
Db 817 AAAGAGAGAGAGAGAGAA 834

RESULT 4
LOCUS B0421827 880 bp mRNA linear EST 23-MAY-2002
DEFINITION AGENCOURT_7801623 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6050568
5', mRNA sequence.
ACCESSION B0421827

```

```

VERSION      B0421827.1  GI:21117142
KEYWORDS
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 880)
              NIH-MGC http://mgc.ncl.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE        Unpublished (1999)
JOURNAL      Contact: Robert Strausberg, Ph.D.
              Email: c9apbs-remail.nih.gov
              Tissue Procurement: ATCC/DCTD/DTF
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNLN at:
              http://lmage.llnl.gov
              Plate: LLM11303 row: a column: 01
              High quality sequence stop: 580.
FEATURES
source       1..880
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:6030568"
              /clone_11b="NIH_MGC_72"
              /tissue_type="melanotic melanoma"
              /lab_host="DH10B (phage-resistant)"
              /note="Organ: skin; Vector: pCMV-SPOrt6; Site_1: NCL1;
              Site_2: Salt; Cloned unidirectionally; Primer: Oligo dT.
              Average insert size 2 kb. Library constructed by Life
              Technologies."
BASE COUNT   238 a 208 c 256 g 176 t 2 others
ORIGIN
Query Match      87.5%; Score 705.4; DB 14; Length 880;
Best Local Similarity 96.8%; Pred. No. 1.8e-114;
Matches 760; Conservative 1; Mismatches 19; Indels 5; Gaps 4;
OY 11 CTGGGCGCAAAATGAAATGACGCGCTGCGCAGCGCTTGCGCCAGCGCGCGCCCTGG 70
Db 9 CTGGGCGCAAAATGAGATGACGCGCTTGCGCAGCGCTTGCGCCAGCGCGCGCCCTGG 68
OY 71 GACCAAGGTGAGCAACCCCGTACCTAAARATGAAAGCGTGGGCTTGCTGCCCTGG 130
Db 69 GAGCAGAGGTGAGCGACGCCCATACCTAAAGATGAAAGCGTGGGCTTGCTGCCCTGG 128
OY 131 CTTGCGGGGGCCCTGCTGGGAAACCGCGCTGGAGAGCCAGATCTCCACTGTGA 190
Db 129 CTTGCGGGGGCCCTGCTGGGAAACCGCGCTGGAGAGCCAGATCTCCACTGTGA 188
OY 191 GCATCAGAGGCTCTGTGATGATAGTAAATGGAAATGGCCAGGTGACCCCAAGAAG 250
Db 189 GCATCAGAGGCTCTGTGATGATAGTAAATGGAAATGGCCAGGTGACCCCAAGAAG 248
OY 251 ACCATTCAGATGGGATCTTTCGGATCAATCCAGATGGCAGCCAGTGTGAGAGTG 310
Db 249 ACCATTCAGATGGGATCTTTCGGATCAATCCAGATGGCAGCCAGTGTGAGAGTG 308
OY 311 CTTATGCGCGCTGAGAGGCCCACTCAGAGAGTGTGGAGAGATATGTACCGGATG 370
Db 309 CTTATGCGCGCTGAGAGGCCCACTCAGAGAGTGTGGAGAGATATGTACCGGATG 368
OY 371 AAGGATATGGGGAACAGATGATCTTTCACACCATTCGAGAGAACTAGTACGTAGTG 430
Db 369 AAGGATATGGGGAACAGATGATCTTTCACACCATTCGAGAGAACTAGTACGTAGTG 428
OY 431 GGCCGGAATGGAGATCCAGTGAAGTGAAGTGAAGGATCCGATGAGTCAAGATAT 490
Db 429 GGCCGGAATGGAGATCCAGTGAAGTGAAGTGAAGGATCCGATGAGTCAAGATAT 488
OY 491 AGCGCACCCTCAATTTGGCTGAGAGCAATGTGGAGGAATAGAGATCAATCTATT 550

```

```

Db 489 AGCGCACCCTCAATTTGGCTGAGAGCAATGTGGAGGAATAGAGATCAATCTATT 548
OY 551 GAATCTTTCCCGAGAGAGCTGACATATGTTAAAGCAATTTGCGATGAGCAACAGAT 610
Db 549 GAATCTTTCCCGAGAGAGCTGACATATGTTAAAGCAATTTGCGATGAGCAACAGAT 608
OY 611 CTTTGTGACCAATGCCCTGACATATGCGATGATGAGATGAGAACCACTGAGCAGCCAC 670
Db 609 CTTTGTGACCAATGCCCTGACATATGCGATGATGAGATGAGAACCACTGAGCAGCCAC 668
OY 671 ACT-GGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 728
Db 669 ACTGGGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 728
OY 729 -ATGTTTACTGAAATTAATGCT--AAAAATATGAAACCAAGTAAACCAACCAACCA 785
Db 729 AATGTTTACTGAAATTAATGCTGAAAAAATATGAAACCAACCAACCAACCAACCA 788
OY 786 AAGA 790
Db 789 AAAAA 793
RESULT 5
LOCUS      B1763950 875 bp mRNA linear EST 25-SEP-2001
DEFINITION 603049829P1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5189862 5',
            mRNA sequence.
ACCESSION  B1763950
VERSION     B1763950.1 GI:15755528
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 875)
              NIH-MGC http://mgc.ncl.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE        Unpublished (1999)
JOURNAL      Contact: Robert Strausberg, Ph.D.
              Email: c9apbs-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNLN at:
              http://lmage.llnl.gov
              Plate: LLM11474 row: j column: 07
              High quality sequence start: 3
              High quality sequence stop: 834.
FEATURES
source       1..875
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:5189862"
              /clone_11b="NIH_MGC_116"
              /lab_host="DH10B"
              /note="Organ: pooled colon, kidney, stomach; Vector:
              pCMV-SPOrt6; Site_1: NCL1; Site_2: EcoRV (destroyed); RNA
              source anonymous pool of 3 colons, age 26 yo male, 49 yo
              female, 71 yo male colon; 46 yo male kidney, and pool of 2
              stomachs, 62 yo male and 70 yo female. Library is
              oligo-dT primed and directionally cloned (EcoRV site is
              destroyed upon cloning). Average insert size 1.4 kb,
              insert size range 1-3 kb. Library is normalized and
              enriched for full-length clones and was constructed by C.
              Gruber (Invitrogen). Research Genetics tracking code
              023. Note: This is a NIH-MGC Library."
BASE COUNT   226 a 209 c 259 g 181 t
ORIGIN

```



Query Match 86.7%; Score 698.6; DB 13; Length 875;  
 Best Local Similarity 97.3%; Pred. No. 2.7e-113;  
 Matches 762; Conservative 1; Mismatches 15; Indels 5; Gaps 5;

11 CTGGGGCCAAAGTGAAGTCCAGGCTGCTGCCAGCCCTGGGGCCAGGGGGCCCTGG 70  
 92 CTGGGGCCAAAGTGAAGTCCAGGCTGCTGCCAGCCCTGGGGCCAGGGGGCCCTGG 151  
 71 GACCAAAAGTGAAGTGAAGTCCAGGCTGCTGCCAGCCCTGGGGCCAGGGGGCCCTGG 130  
 152 GACCAAAAGTGAAGTGAAGTCCAGGCTGCTGCCAGCCCTGGGGCCAGGGGGCCCTGG 211  
 131 CTCTGGGGGGCCCTGGGGGGAGCCCTGGGGCCCTGGGGCCAGGGGGCCCTGG 190  
 212 CTCTGGGGGGCCCTGGGGGGAGCCCTGGGGCCCTGGGGCCAGGGGGCCCTGG 271  
 191 GCATGAGAGGCTCTGGTGGATGAATAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 250  
 272 GCATGAGAGGCTCTGGTGGATGAATAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 331  
 251 ACCATTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 310  
 332 ACCATTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 391  
 311 CCTATGCCCCGCTCAGAGGCCCCCCTCAGAGGCCCCCCTCAGAGGCCCCCCTCAGAGG 370  
 392 CCTATGCCCCGCTCAGAGGCCCCCCTCAGAGGCCCCCCTCAGAGGCCCCCCTCAGAGG 451  
 371 AAGAGTATGGGGAACAGATGATCTTCCAGCCATCCAGAGTGAAGTGAAGTGAAGTGAAG 430  
 452 AAGAGTATGGGGAACAGATGATCTTCCAGCCATCCAGAGTGAAGTGAAGTGAAGTGAAG 511  
 431 GGGCGAATGAGAAATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 490  
 512 GGGCGAATGAGAAATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 571  
 491 AGCGGACCCCTCAAGTTGCGTGTGAGAGCATTTGAGAGAAATGAGAGTGAAGTGAAGTGAAG 550  
 572 AGCGGACCCCTCAAGTTGCGTGTGAGAGCATTTGAGAGAAATGAGAGTGAAGTGAAGTGAAG 631  
 551 GAATTTCTTTCCGAGAGGCTGAGAAATGTTAAAGACAAC-TTTCAGTGAAGCAGACAGA 609  
 632 GAATTTCTTTCCGAGAGGCTGAGAAATGTTAAAGACAAC-TTTCAGTGAAGCAGACAGA 691  
 610 TCTTTGTGACCATGCTGACATAT-TGCGATGATGAGTATGAAACACT-GGAGCAGCC 667  
 692 TCTTTGTGACCATGCTGACATAT-TGCGATGATGAGTATGAAACACT-GGAGCAGCC 751  
 668 CACACTGGCTTGAATGATCAACCCAGAGGGGAAATGTTGGCAATGCC-TTTTATATA 726  
 752 CACACTGGCTTGAATGATCAACCCAGAGGGGAAATGTTGGCAATGCC-TTTTATATA 811  
 727 TTAATGTTTACTGAAA-TTAATGAAAATAATGAAAACCAAAAGTAAAAA 785  
 812 TTAATGTTTACTGAAA-TTAATGAAAATAATGAAAACCAAAAGTAAAAA 871  
 786 AAA 788  
 872 AAA 874

RESULT 6  
 BI550162 773 bp mRNA linear EST 05-SRP-2001  
 LOCUS 603193045F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:5264339 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI550162  
 VERSION BI550162.1 GI:15437474  
 KEYWORDS EST  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 773)  
 AUTHORS NIH-MGC http://mgi.mgi.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: rstraub@mail.nih.gov  
 Tissue procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA library preparation: Michael J. Brownstein (NMGRI), Shiroki  
 Toshiyuki and Piero Carninci (RIKEN)  
 DNA library arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.lnl.gov  
 Plate: LAM1166 row: 1 column: 12  
 High quality sequence stop: 765.  
 Location/Qualifiers  
 1. 773  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5264339"  
 /clone\_lib="NIH\_MGC\_95"  
 /issue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pBluescript (modified  
 pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',  
 size-selected for average insert size 2.5 kb and  
 normalized to 10^6. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NMGRI/NHRI, National  
 Institutes of Health). Note: this is a NIH-MGC Library."  
 BASE COUNT 204 a 178 c 226 g 163 t 2 others  
 ORIGIN

Query Match 86.0%; Score 693; DB 13; Length 773;  
 Best Local Similarity 96.4%; Pred. No. 2.8e-112;  
 Matches 740; Conservative 1; Mismatches 22; Indels 5; Gaps 3;

21 AGTGAAGTCCAGGCTGCTGCCAGCCCTGGGGCCAGGGGGCCCTGGAGCAAGT 80  
 1 AGCGGGGCTCAGCGGCTTCCAGCGCTTGGGCGAGCGGGCCCTGGAGCAAGT 60  
 81 GGAGCAACCCCTTACCCCTAAATGAAGAGGCTGGGCTGGCTGGCTGGCTGGG 140  
 61 GGAGCAACCCCTTACCCCTAAATGAAGAGGCTGGGCTGGCTGGCTGGG 120  
 141 CCTGCTGGGAGCGGCTGGGCTGGGAGAGCCAGATCTCCACTGTGAGATGCAAGG 200  
 121 CCTGCTGGGAGCGGCTGGGCTGGGAGAGCCAGATCTCCACTGTGAGATGCAAGG 180  
 201 CTCTGGTGAATGAATGAATGGGAATGGCCAGGTGAGACCCCAAGAACATTGACA 260  
 181 CTCTGGTGAATGAATGAATGGGAATGGCCAGGTGAGACCCCAAGAACATTGACA 240  
 261 TGGGATCTTCCGATCAATCCAGATGAGCAGCAGTCACTGGTGGAGGCTTATGCC 320  
 241 TGGGATCTTCCGATCAATCCAGATGAGCAGCAGTCACTGGTGGAGGCTTATGCC 300  
 321 GCTCAGAGGCCACCTCAGACAGCTGCTGGAGAGATATGTACCGGATGAAGAGTATG 380  
 301 GCTCAGAGGCCACCTCAGACAGCTGCTGGAGAGATATGTACCGGATGAAGAGTATG 360  
 381 GGAACAGATGATCTTCCAGCAGCTGCGAAGAACTAGTATGATGGCCGGAATG 440  
 361 GGAACAGATGATCTTCCAGCAGCTGCGAAGAACTAGTATGATGGCCGGAATG 420  
 441 GGAATCCAGTGAATGAGTGAAGAGCATCCAGATGAGATGAGATGAGATGAGATGAG 500  
 421 GGAATCCAGTGAATGAGTGAAGAGCATCCAGATGAGATGAGATGAGATGAGATGAG 480  
 501 TCAAGTTTGGCGTGAAGAGCATTTGAGAGAAATGAGAGTGAAGTGAATTTCTTTT 560

```

Db 481 TCAAGTTGGCTGTAGACATTTGGAGAAATACGAGATACCAATGTAATCTTTT 540
Oy 561 CCCGAGAGCTGACAAATGTTAAAGACAACTTGCAGTAAGCAACAGATC-TTTGTGAC 619
Db 541 CCCGAGAGCTGACAAATGTTAAAGACAACTTGCAGTAAGCAACAGATCTTTTGTGAC 600
Oy 620 CATGCCCTGACATATGCGATGATGAGCTATGACCACTGAGACGCCACACTGCTTG 679
Db 601 CATGCCCTGACATATGCGATGATGAGCTATGACCACTGAGACGCCACACTGCTTG 660
Oy 680 ATGGATACCCCGAGAGGGG--AAAATGGTGCAATGCTTTTATATTTATGTTTAA 737
Db 661 ATGGATACCCCGAGAGGGGAAATGTTGGCAATGCTTTTATATATGTTTAA 720
Oy 738 CTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 785
Db 721 CTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 766

```

RESULT 7  
 BG824549 742 bp mRNA linear EST 22-MAY-2001  
 LOCUS 602728454P1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4868100 5',  
 DEFINITION mRNA sequence.

ACCESSION BG824549  
 VERSION BG824549.1 GI:14172136  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 742)  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: NIH Intramural Sequencing Center  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1CM1737 row: k column: 13  
 High quality sequence stop: 740.

FEATURES  
 Source Location/Qualifiers  
 1..742

Source  
 1..742  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4868100"  
 /clone\_11b="NIH\_MGC\_15"  
 /tissue\_type="adenocarcinoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: colon; Vector: pOTB1; site\_1: XhoI; site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)"  
 BASE COUNT 180 a 184 c 233 g 144 t 1 others  
 ORIGIN

Query Match 83.9%; Score 676.6; DB 12; Length 742;  
 Best Local Similarity 98.3%; Pred. No. 2.2e-109;  
 Matches 682; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

Oy 11 CTGGGGCCAAAGTGAAGCCAGCGGTGCGCAGCGCTTGGCCAGCGGGGGCCCTCG 70  
 Db 48 CTGGGGCCAAAGTGAAGTCAAGCGGTCTCCAGCGCTTGGCCAGCGGGGGCCCTCG 107

```

Oy 71 GACCAAGGTGGAGCAACCCCGTTACCTTAARATGAAAAGGTGGGTTGGTGGCCCTG 130
Db 108 GAGCAGAGTGGAGCGACCCCAATACGTTAAAGATGAAGGGCTTGGTGGCCCTG 167
Oy 131 CTCTGGGGGGCCCTGCTGGGAACCGCTTGGCTCGAGAGACCGAGATCTCCACTGTGA 190
Db 168 CTCTGGGGGGCCCTGCTGGGAACCGCTTGGCTCGAGAGACCGAGATCTCCACTGTGA 227
Oy 191 GCATGAGGGCTCTGGTGGATGACATGAGTGAATGGAAATGGCCAGGTGACCCCAAGAG 250
Db 228 GCATGAGGGCTCTGGTGGATGACATGAGTGAATGGAAATGGCCAGGTGACCCCAAGAG 287
Oy 251 ACCATTGAGATGGATCTTCCGGATCAATCCAGATGGCAGCCAGTCACTGGTGGAGTG 310
Db 288 ACCATTGAGATGGATCTTCCGGATCAATCCAGATGGCAGCCAGTCACTGGTGGAGTG 347
Oy 311 CCTTATGCCCCGCTCAGAGGCCCACTCAGAGAGCTCTGGAGAGATATGTGACCGGATG 370
Db 348 CCTTATGCCCCGCTCAGAGGCCCACTCAGAGAGCTCTGGAGAGATATGTGACCGGATG 407
Oy 371 AAGAGATATGGGGAACAGATGATCCTTCCACCATCGCAAGAACTAGCTAGCTGATG 430
Db 408 AAGAGATATGGGGAACAGATGATCCTTCCACCATCGCAAGAACTAGCTAGCTGATG 467
Oy 431 GGCCGGAATGAGAAATCCAGTGAATGAGCTGACCTGCAAGGCACTCGAATGACTCAGATAT 490
Db 468 GGCCGGAATGAGAAATCCAGTGAATGAGCTGACCTGCAAGGCACTCGAATGACTCAGATAT 527
Oy 491 AGCGGACCCCTCAAGTTTCCGCTGTGAGACCATTTGGAGAGATAGAGATGAATCACTATT 550
Db 528 AGCGGACCCCTCAAGTTTCCGCTGTGAGACCATTTGGAGAGATAGAGATGAATCACTATT 587
Oy 551 GAATCTTTTCCGAGAGAGCTGACATATGTTAAAGACAACTTGGAGTGAAGCAAGAT 610
Db 588 GAATCTTTTCCGAGAGAGCTGACATATGTTAAAGACAACTTGGAGTGAAGCAAGAT 647
Oy 611 CTCTGACCAATGCCCTGACATATGATGATGATGATGATGATGATGATGATGATGATG 670
Db 648 CTCTGACCAATGCCCTGACATATGATGATGATGATGATGATGATGATGATGATGATG 707
Oy 671 ACTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 704
Db 708 ACTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 741

```

RESULT 8  
 BF344334 697 bp mRNA linear EST 22-NOV-2000  
 LOCUS 602014708P1 NCI\_CGAP\_Brn64 Homo sapiens cDNA clone IMAGE:4150512  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BF344334  
 VERSION BF344334.1 GI:11291554  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 697)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: David N. Louis, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM9413 row: p column: 01  
 High quality sequence stop: 692.

FEATURES  
source  
Location/Qualifiers  
1. 697  
/organism="Homo sapiens"  
/db.xref="taxon:9606"  
/clone="IMAGE:4150512"  
/clone\_lib="NCI\_CGAP\_Brn64"  
/lab\_type="fibroblastoma with EGFR amplification"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: brain; Vector: pCMV-Sport6; Site: 1; NotI;  
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.57 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 173 a 167 c 220 g 137 t  
ORIGIN

Query Match 83.7%; Score 674.4; DB 12; Length 697;  
Best Local Similarity 98.7%; Pred. No. 5.4e-109;  
Matches 689; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

13 GGGGCCAAAGTGAAGTCCAGCGGTCTGCGACGCTTGGGCCACGCGCGCCCTGGGA 72  
1 GGGGCCAAAGTGAAGTCCAGCGGTCTTCCAGCGCTTGGGCCACGCGCGCCCTGGGA 60  
73 CCAAGGTGAGACACCCGCTTAAATGAAGGCTGGGCTGGCTGGCTGGCTGGCT 132  
61 GCAGAGGTGAGACACCCGCTTAAAGTGAAGGCTGGGCTGGCTGGCTGGCT 120  
133 TCTGGGGGCGCTGCGGAAACCGCTGGGCTGGAGAGCGAGATCTCCACTGTGGAGC 192  
121 TCTGGGGGCGCTGCTGGGAAACCGCTGGGCTGGAGAGCGAGATCTCCACTGTGGAGC 180  
193 ATGCAGGGCTGTGTGATGAATGGAATTTGCCAGGTGAGACCCCAAGAGAGC 252  
181 ATGCAGGGCTGTGTGATGAATGGAATTTGCCAGGTGAGACCCCAAGAGAGC 240  
253 CATTCAGATGGATCTTCCGATCAATCCATGCGACCGCTGAGTGGAGGTGCC 312  
241 CATTCAGATGGATCTTCCGATCAATCCATGCGACCGCTGAGTGGAGGTGCC 300  
313 TTATGCGGCTGAGAGGCGCCACCTCAAGAGCTGTGAGAGATATGATGAGCGATGAA 372  
301 TTATGCGGCTGAGAGGCGCCACCTCAAGAGCTGTGAGAGATATGATGAGCGATGAA 360  
373 GGAGTATGGGGAACAGATTTGATCTTCCACCCATCGCAAGAACTAGTGTAGTGG 432  
361 GGAGTATGGGGAACAGATTTGATCTTCCACCCATCGCAAGAACTAGTGTAGTGG 420  
433 CGGGAATGGAGATCCAGTGAATGAGTACTAAGGCAATCCCAATCGATGATTTAG 492  
421 CGGGAATGGAGATCCAGTGAATGAGTACTAAGGCAATCCCAATCGATGATTTAG 480  
493 CGGCAACCTCAAGTTTGCCTGTGAGAGCATTTGAGAGAAATGAGAGTCAATTA 552  
481 CGGCAACCTCAAGTTTGCCTGTGAGAGCATTTGAGAGAAATGAGAGTCAATTA 540  
553 ATTTCTTTTCCGAGAGGCTGACATTTTAAAGACAACTTTTACAGTAAGCAAGATCT 612  
541 ATTTCTTTTCCGAGAGGCTGACATTTTAAAGACAACTTTTACAGTAAGCAAGATCT 600  
613 TTGTGACCATTCGCTGACATATTCGATGATGATGATGATGATGATGATGATGAT 672  
601 TTGTGACCATTCGCTGACATATTCGATGATGATGATGATGATGATGATGATGAT 659  
673 TGGCTTGTATGATCAACCCCGAGAGGGGAAATGGTGG 710  
660 TGGCTTGTATGATCAACCCCGAGAGGGGAAATGGTGG 697

RESULT 9  
BI762759 887 bp mRNA linear EST 25-SEP-2001  
LOCUS BI762759  
DEFINITION 603048045F1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5188307 5',  
mRNA sequence.

ACCESSION BI762759  
VERSION BI762759.1 GI:15754325  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 887)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LML at:  
http://image.llnl.gov  
Plate: LML11470 row: 1 column: 12  
High quality sequence stop: 775.

FEATURES  
source  
Location/Qualifiers  
1. 887  
/organism="Homo sapiens"  
/db.xref="taxon:9606"  
/clone="IMAGE:5188307"  
/clone\_lib="NIH\_MGC\_116"  
/lab\_host="DH10B"  
/note="Organ: pooled colon, kidney, stomach; Vector:  
pCMV-Sport6; Site: 1; NotI; Site: 2; EcoRV (destroyed); RNA  
source anonymous pool of 3 colons, age 26 yo male, 49 yo  
female, 71 yo male colon; 46 yo male kidney, and pool of 2  
stomachs, 62 yo male and 70 yo female. Library is  
oligo-dT primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.4 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
023. Note: this is a NIH\_MGC Library."

BASE COUNT 235 a 228 c 252 g 172 t  
ORIGIN

Query Match 83.3%; Score 671.4; DB 13; Length 887;  
Best Local Similarity 95.6%; Pred. No. 1.6e-108;  
Matches 743; Conservative 1; Mismatches 27; Indels 6; Gaps 5;

9 GCGTGGGGCCAAAGTGAAGTCCAGCGGTCTGCGACGCTTGGGCCACGCGCGCCCT 68  
60 GACTGGGGCCAAAGTGAAGTCCAGCGGTCTTCCAGCGCTTGGGCCACGCGCGCCCT 119  
69 GGGACCAAGGTGGAGCAACCCCGTTACCTTAAATGAAGGCTGGGTTGGCTGGCC 128  
120 GGGAGCAGAGGTGGAGCGACCCCATTTAGCTTAAATGAAGGCTGGGTTGGCTGGCC 179  
129 TGCCTTCTGGGGCCCTGCTGCGAGACCGCTGGGCTCGGAGCAGCCAGATCTCCACTGG 188  
180 TGCCTTCTGGGGCCCTGCTGCGAGACCGCTGGGCTCGGAGCAGCCAGATCTCCACTGG 239  
189 GAGCATGCAAGGCTGTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 248  
240 GAGCATGCAAGGCTGTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 299  
249 AGACCATTCAGATGGGATCTTCCGAGTCAATCCAGATGAGCAGCAGTCAATGAGGAG 308  
300 AGACCATTCAGATGGGATCTTCCGAGTCAATCCAGATGAGCAGCAGTCAATGAGGAG 359  
309 TGCCTTATGCCCGCTCAGAGGCCACCTCACAGACTGCTGGAGAGATATGACCGGA 368  
360 TGCCTTATGCCCGCTCAGAGGCCACCTCACAGACTGCTGGAGAGATATGACCGGA 419  
369 TGAAGAGTATGGGAGACAGATTTGATCTTCCACCATTCGCAAGAACTAGTACGTGAG 428

Db 420 TGAAGAGTATGGGGAACAGATGATCCTTCACCCATCGCAAGAACTACCTGACCTGAG 479  
OY 429 TGGGCCGGAATGAGAAATCCAGTGAATGAGCATCAAGCAATCCGAATCGACTCAGATA 488  
Db 480 TGGGCCGGAATGAGAAATCCAGTGAATGAGCATCAAGCAATCCGAATCGACTCAGATA 539  
OY 489 TTAAGGGGCAACCTCAAGTTTGGCTGTGAGAGCATTTGTGAGGAATACGAGATGAACTCA 548  
Db 540 TTAAGGGGCAACCTCAAGTTTGGCTGTGAGAGCA-TGTGAGGAAATACGAGGTGAACTCA 598  
OY 549 TTAAGTTCTTTCCCGAGAGGCTGACATGTTAAAGCAAACTTGGAGTAAGCAAGACAG 608  
Db 559 TTGAATTTCTTTCCCGAGAGGCTGACATGTTAAAGCAAACTTGGAGTAAGCAAGACAG 658  
OY 609 ATCTTTTGACCATGCGCTGACATATGATGATGAGTATGAAACCACTGGAGCAGCCC 668  
Db 659 ATCTTTTGACCATGCGCTGACATATGATGATGAGTATGAAACCACTGGAGCAG-CC 717  
OY 669 ACACGGCTTATGATGATCAACCCCGAGAGGGGAAATGG-TGGCAATGCTTTTATA-TA 726  
Db 718 ACACGGCTTATGATGATCAACCCCGAGAGGGGAAATGGTGGCAATGCTTTTATACTA 777  
OY 727 TTAATGTTTACT--GAATTAAGTAAATAATATGAACCAAAAGTAAAAA 781  
Db 778 TTAATGTTTACTACATTAATGAGTGGACACCTATGAACCATTAAGTAAAAAACA 834

RESULT 10  
B0668530 673 bp mRNA linear EST 15-JUL-2002  
LOCUS AGENCOURT\_8211118 NIH\_MGC\_102 Homo sapiens cDNA clone IMAGE:6276413  
DEFINITION 5', mRNA sequence.  
ACCESSION B0668530  
VERSION B0668530.1 GI:2178777  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 673)  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapds-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LCM2461 row: c column: 06  
High quality sequence start: 208  
High quality sequence stop: 350.  
Location/Qualifiers

FEATURES  
Source  
1..673  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6276413"  
/clone\_id="NIH\_MGC\_102"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: salivary gland; Vector: pOTB7; site\_1: XhoI;  
site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAGAG(C) library constructed  
by Ling Hong in the laboratory of Gerald M. Rubin  
(University of California, Berkeley) using ZAP-cDNA  
synthesis kit (Stratagene) and Superscript II RT (Life  
Technologies). Note: this is a NIH\_MGC Library."  
BASE COUNT 196 a 149 c 184 g 144 t  
ORIGIN

Query Match 83.1%; Score 669.8; DB 14; Length 673;  
Best Local Similarity 99.7%; Pred. No. 3.5e-108;  
Matches 671; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 118 TTGGCTGGCCCTTCTTCTTGGGGGCCCTGCTGGGAACCCGCTGGCTCGGAGAGCCAGCA 177  
Db 1 TTGGCTGGCCCTTCTTCTTGGGGGCCCTGCTGGGAACCCGCTGGCTCGGAGAGCCAGCA 60  
OY 178 TCTCCACTGTGAGCATCAGAGGCTCTGTGGATGACTAGAAATGGAAATGGCCAGGT 237  
Db 61 TCTCCACTGTGAGCATCAGAGGCTCTGTGGATGACTAGAAATGGAAATGGCCAGGT 120  
OY 238 GGACCCCAAGAAAGACCATTCAGATGGATCTTTCCGATCAATCCAGATGGCAGCAGT 297  
Db 121 GGACCCCAAGAAAGACCATTCAGATGGATCTTTCCGATCAATCCAGATGGCAGCAGT 180  
OY 298 AGTGTGAGAGTGGCTTTATGCCCGCTCAGAGGCCCACTCAGAGCTGCTGGAGAGAT 357  
Db 181 AGTGTGAGAGTGGCTTTATGCCCGCTCAGAGGCCCACTCAGAGCTGCTGGAGAGAT 240  
OY 358 ATGTGACCGGATGAGAGAGTATGGGGAACAGATGATCTTCCACCATCGAAGACTA 417  
Db 241 ATGTGACCGGATGAGAGAGTATGGGGAACAGATGATCTTCCACCATCGAAGACTA 300  
OY 418 CGTACGTTAGTGGCGCGGAATGGAAATCCAGTGAACCTGACCTCAAGCAGCATCCGAT 477  
Db 301 CGTACGTTAGTGGCGCGGAATGGAAATCCAGTGAACCTGACCTCAAGCAGCATCCGAT 360  
OY 478 CGACTCAATATTTAGCCGACCCCTCAAGTTTCCGTGTAGAGCATTTGGAGAAATACGA 537  
Db 361 CGACTCAATATTTAGCCGACCCCTCAAGTTTCCGTGTAGAGCATTTGGAGAAATACGA 420  
OY 538 GGATGAACTCATTTGATCTTTTCCCGAGAGCTGACATGTTAAAGCAAACTTTCAG 597  
Db 421 GGATGAACTCATTTGATCTTTTCCCGAGAGCTGACATGTTAAAGCAAACTTTCAG 480  
OY 598 TAAAGCAACAGATCTTTGTGACCATGCCCTGCATATGCGATGATGATGAACAC 657  
Db 481 TAAAGCAACAGATCTTTGTGACCATGCCCTGCATATGCGATGATGATGAACAC 540  
OY 658 TGGAGCAGCCCAACATGCGCTGTGATGATGATGATGATGATGATGATGATGATGAT 717  
Db 541 TGGAGCAGCCCAACATGCGCTGTGATGATGATGATGATGATGATGATGATGATGAT 600  
OY 718 TTTTATATATATGTTTCTTACTGAAATTAAGTAAATTAAGTAAATTAAGTAAATTAAG 777  
Db 601 TTTTATATATATGTTTCTTACTGAAATTAAGTAAATTAAGTAAATTAAGTAAATTAAG 660  
OY 778 AAAAAAAAAAAGA 790  
Db 661 AAAAAAAAAAAGA 673  
RESULT 11  
B1909759 720 bp mRNA linear EST 16-OCT-2001  
LOCUS B1909759  
DEFINITION 60307084F1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5219603 5',  
mRNA sequence.  
ACCESSION B1909759  
VERSION B1909759.1 GI:16173041  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 720)  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapds-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LLM1552 row: a column: 12  
 High quality sequence stop: 717.

FEATURES  
 source location/Qualifiers  
 1..720

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5219603"  
 /clone\_1lb="NIH\_MGC\_118"  
 /tissue\_type="Leukocyte"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV  
 (destroyed); RNA source leukocytes from anonymous pool of  
 non-activated adult donors. Library is oligo-dT primed  
 and directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.7 kb, insert size range  
 1.2-3.3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 027. Note:  
 this is a NIH\_MGC Library."

BASE COUNT 176 a 175 c 230 g 139 t  
 ORIGIN

Query Match 82.0%; Score 660.8; DB 13; Length 720;  
 Best Local Similarity 97.8%; Pred. No. 1.3e-106;

Matches 679; Conservative 1; Mismatches 13; Indels 1; Gaps 1;

QY 11 CTGGGGCCAAAGTGAAGTCCAGGCGTCTGCCAGGCTTGGGCCACGGCGGCGCCCTGG 70  
 DB 27 CTGGGGCCAAAGTGAAGTCCAGGCGTCTGCCAGGCTTGGGCCACGGCGGCGCCCTGG 86  
 QY 71 GACCAAGGTGAGACMACCCGTTACCTTAARATGAAGGCTGGGCTGGCTGGCCCTG 130  
 DB 87 GAGCAGAGGTGAGCAGCACCCTATACGTTAAGATGAAGAGCTGGGCTGGCTGGCCCTG 146  
 QY 131 CTTCCGGGGGCGCTGCTGGGGAACCGGCTGGGCTGGGAGACCCAGAGATCTCCACTGTGA 190  
 DB 147 CTTCTGGGGGCGCTGCTGGGGAACCGGCTGGGCTGGGAGACCCAGAGATCTCCACTGTGA 206  
 QY 191 GCATGCAAGGCGCTGCTGGGGAACCGGCTGGGCTGGGAGACCCAGAGATCTCCACTGTGA 250  
 DB 207 GCATGCAAGGCGCTGCTGGGGAACCGGCTGGGCTGGGAGACCCAGAGATCTCCACTGTGA 266  
 QY 251 ACCATTGAGATGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCACTGGTGGAGTGTG 310  
 DB 267 ACCATTGAGATGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCACTGGTGGAGTGTG 326  
 QY 311 CCTTATCCCGGCTGAGAGGCGCCACCTCAAGAGTGTGTCGAGAGATATGTGACCGGATG 370  
 DB 327 CCTTATCCCGGCTGAGAGGCGCCACCTCAAGAGTGTGTCGAGAGATATGTGACCGGATG 386  
 QY 371 AAGGAGTATGGGGAACAGATGATCTTCCACCATCGCAAGAACTGTCGATACGCTGTGTG 430  
 DB 387 AAGGAGTATGGGGAACAGATGATCTTCCACCATCGCAAGAACTGTCGATACGCTGTGTG 446  
 QY 431 GGCCGGAATGGGAATTCATGTAATGAGTCACTCAAGAGATCCGAATCGACTGATATAT 490  
 DB 447 GGCCGGAATGGGAATTCATGTAATGAGTCACTCAAGAGATCCGAATCGACTGATATAT 506  
 QY 491 AGCGGACCCCTCAAGTTGGTGTGAGAGATTTGGAGAAATACAGAGATGAACATCAT 550  
 DB 507 AGCGGACCCCTCAAGTTGGTGTGAGAGATTTGGAGAAATACAGAGATGAACATCAT 566  
 QY 551 GAATTTTTCCTCCGAGAGGCTGACAAATGTTAAAGACAAATTTGACAGTAAGCAAGAT 610  
 DB 567 GAATTTTTCCTCCGAGAGGCTGACAAATGTTAAAGACAAATTTGACAGTAAGCAAGAT 626  
 QY 611 CTTTGTGACCATCCCTGAC -ATATCGATGATGACTATGACCACTGAGACAGCCCA 669

DB 627 CTTGCGACCATGGCTGACACATATCGCATATGACTATGATACCATGAGAGGCCA 686  
 QY 670 CACTGGCTGATGATGATACACCCAGAGAGGAGAA 703  
 DB 687 CACTGGCTGATGATGATACACCCAGAGAGGAGACA 720

RESULT 12

LOCUS BE799495 914 bp mRNA linear EST 20-SEP-2000  
 DEFINITION 601589289F1 NIH\_MGC\_7 Homo sapiens CDNA clone IMAGE:3943223 5',  
 mRNA sequence.  
 ACCESSION BE799495  
 VERSION BE799495.1 GI:10220693  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 914)  
 NIH-MGC <http://mgc.ncl.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 CONTACT Robert Strausberg, Ph.D.  
 COMMENT Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at: [image.llnl.gov](http://image.llnl.gov)  
 Plate: LICW798 row: b column: 24  
 High quality sequence start: 4  
 High quality sequence stop: 788.

FEATURES  
 source location/Qualifiers

1..914  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3943223"  
 /clone\_1lb="NIH\_MGC\_7"  
 /tissue\_type="small cell carcinoma"  
 /cell\_line="MGC3"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: Lung; Vector: pOT8; Site.1: XhoI; Site.2:  
 EcoRI; CDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGACGAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the Laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 230 a 216 c 292 g 176 t  
 ORIGIN

Query Match 80.7%; Score 650.8; DB 12; Length 914;  
 Best Local Similarity 92.9%; Pred. No. 6.5e-105;  
 Matches 736; Conservative 1; Mismatches 48; Indels 7; Gaps 5;

QY 11 CTGGGGCCAAAGTGAAGTCCAGGCGTCTGCCAGGCTTGGGCCACGGCGGCGCCCTGG 70  
 DB 127 CTGGGGCCAAAGTGAAGTCCAGGCGTCTGCCAGGCTTGGGCCACGGCGGCGCCCTGG 186  
 QY 71 GACCAAGGTGAGACMACCCGTTACCTTAARATGAAGGCTGGGCTGGCTGGCCCTG 130  
 DB 187 GAGCAGAGGTGAGCAGCACCCTATACGTTAAGATGAAGAGCTGGGCTGGCTGGCCCTG 246  
 QY 131 CTTTGGGGGCGCTGCTGGGGAACCGGCTGGGCTGGGAGACCCAGAGATCTCCACTGTGA 190  
 DB 247 CTTTGGGGGCGCTGCTGGGGAACCGGCTGGGCTGGGAGACCCAGAGATCTCCACTGTGA 306  
 QY 191 GCATGCAAGGCGCTGCTGGGGAACCGGCTGGGCTGGGAGACCCAGAGATCTCCACTGTGA 250

Db 307 GCATGACGAGGCTCTGTGGATGATGACTAGATGGAAATTTGCCAGGTGACCCCAAGAAG 366  
 Oy 251 ACCATTGAGATGGATCTTTCCGATCAATCCAGATGGCAGCAGTCACTGGTGGAGTG 310  
 Db 367 ACCATTGAGATGGATCTTTCCGATCAATCCAGATGGCAGCAGTCACTGGTGGAGTG 426  
 Oy 311 CCTTATGCCCCCTCAGAGGCCCTCAGAGAGCTGCTGGAGAGATATGTGACCGGAGT 370  
 Db 427 CCTTATGCCCCCTCAGAGGCCCTCAGAGAGCTGCTGGAGAGATATGTGACCGGAGT 486  
 Oy 371 AAGGATATGGGAGACAGATGATGCTTCCACCCATCCGAGAGACTAGCTAGTGTAGT 430  
 Db 487 AAGGATATGGGAGACAGATGATGCTTCCACCCATCCGAGAGACTAGCTAGTGTAGT 546  
 Oy 431 GGCCGGAATGAGAAATCCAGTGAAGTGAAGTGAAGGATCCGATCCGATCCAGTATAT 490  
 Db 547 GGCCGGAATGAGAAATCCAGTGAAGTGAAGTGAAGGATCCGATCCGATCCAGTATAT 606  
 Oy 491 AGCGGACCCCTCAAGTTTCCGTTGAGAGCATTTGAGAGATGAGAGATGAGATCACTATT 550  
 Db 607 AGCGGACCCCTCAAGTTTCCGTTGAGAGCATTTGAGAGATGAGAGATGAGATCACTATT 666  
 Oy 551 G-AATTTCTTTTCCGAGAGGCTGACAAATGTAAAGCAAACTTTGAGAGAGAGAGAG 609  
 Db 667 GAAATTTCTTTTCCGAGAGGCTGACAAATGTAAAGCAAACTTTGAGAGAGAGAGAG 725  
 Oy 610 TCTTTGTGACATGCTCTGACATATTCGATGATGATGATGATGATGATGATGATGATG 669  
 Db 726 TCTTTGTGACATGCTCTGACATATTCGATGATGATGATGATGATGATGATGATGATG 784  
 Oy 670 CACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729  
 Db 785 CACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842  
 Oy 730 TCTTTTCTGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789  
 Db 843 -GGTTTACTGAACTTATGAAATTAATGAAATTAATGAAATTAATGAAATTAATGAAAT 900  
 Oy 790 AGAGAGAGAGAG 801  
 Db 901 TGGGGGGGAGAG 912

RESULT 13  
 LOCUS Bg704443 791 bp mRNA linear EST 07-MAY-2001  
 DEFINITION 60268633F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:4821008 5',  
 mRNA sequence.  
 ACCESSION Bg704443  
 VERSION Bg704443.1 GI:13977790  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 791)  
 AUTHORS NIH-MGC http://mgi.ncl.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strusberg, Ph.D.  
 Email: cgrabs-remail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1M10727 row: a column: 09  
 High quality sequence stop: 785.  
 Location/Qualifiers  
 1..791

FEATURES  
SOURCE

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4821008"  
 /clone\_1ib="NIH\_MGC\_95"  
 /tissue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /note="Organ: Brain; Vector: pBluescript (modified  
 pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcag  
 ); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTNN-3',  
 size-selected for average insert size 2.5 kb and  
 normalized to R0.5. This is a primary library enriched  
 for full-length clones and constructed using the  
 cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NHGRI, National  
 Institutes of Health). Note: this is a NIH\_MGC Library."  
 BASE COUNT 187 a 196 c 256 g 152 t  
 ORIGIN

Query Match 80.2%; Score 646.4; DB 12; Length 791;  
 Best Local Similarity 97.1%; Pred. No. 4.1e-104;  
 Matches 668; Conservative 1; Mismatches 17; Indels 2; Gaps 1;

Oy 11 CTGGGGCCAAAGTGAAGTCCAGCGGCTGCGCAGCGCTGGGCGACAGGGGGCCCTGG 70  
 Db 106 CTGGGGCCAAAGTGAAGTCCAGCGGCTGCGCAGCGCTGGGCGACAGGGGGCCCTGG 165  
 Oy 71 GACCAAGGTGAGCAACCCCTTACCTAAATATGAAGGCTGGGCTGGGCTGGGCTGG 130  
 Db 166 GAGCAGAGGTGAGCGACCCCAATACGTAAAGATGAAGGCTGGGCTGGGCTGGGCTGG 225  
 Oy 131 CTTTGGGGGGCCCTGCTGGGAAACCCGCTGGGCTGGGAGAGAGCCAGATGTCCTGTA 190  
 Db 226 CTTTGGGGGGCCCTGCTGGGAAACCCGCTGGGCTGGGAGAGAGCCAGATGTCCTGTA 285  
 Oy 191 GCATGACGAGGCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 250  
 Db 286 GCATGACGAGGCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 345  
 Oy 251 ACCATTGAGATGGATCTTTCCGATCAATCCAGATGGCAGCAGTCACTGGTGGAGTG 310  
 Db 346 ACCATTGAGATGGATCTTTCCGATCAATCCAGATGGCAGCAGTCACTGGTGGAGTG 405  
 Oy 311 CCTTATGCCCCCTCAGAGGCCCTCAGAGAGCTGCTGGAGAGATATGTGACCGGAGT 370  
 Db 406 CCTTATGCCCCCTCAGAGGCCCTCAGAGAGCTGCTGGAGAGATATGTGACCGGAGT 465  
 Oy 371 AAGGATATGGGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 430  
 Db 466 AAGGATATGGGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 525  
 Oy 431 GGCCGGAATGAGAAATCCAGTGAAGTGAAGTGAAGGATCCGATCCGATCCAGTATAT 490  
 Db 526 GGCCGGAATGAGAAATCCAGTGAAGTGAAGTGAAGGATCCGATCCGATCCAGTATAT 585  
 Oy 491 AGCGGACCCCTCAAGTTTCCGTTGAGAGCATTTGAGAGAGATGAGAGATGAGATCACTATT 550  
 Db 586 AGCGGACCCCTCAAGTTTCCGTTGAGAGCATTTGAGAGAGATGAGAGATGAGATCACTATT 645  
 Oy 551 GAATTTCTTTCCGAGAGGCTGACATGTTAAAGCAAACTTTGCTGAAGGAAGACAGAT 610  
 Db 646 GAATTTCTTTCCGAGAGGCTGACATGTTAAAGCAAACTTTGCTGAAGGAAGACAGAT 705  
 Oy 611 CTTTGTGACCATGCTTGCATATGATGATGATGATGATGATGATGATGATGATGATGATG 670  
 Db 706 CTTTGTGACCATGCTTGCATATGATGATGATGATGATGATGATGATGATGATGATGATG 763  
 Oy 671 ACTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 698  
 Db 764 CATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 791

RESULT 14  
BF312281







GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:08:03 ; Search time 142.849 Seconds  
(without alignments)  
2869.213 Million cell updates/sec

Title: US-10-082-502-17

Perfect score: 960  
Sequence: 1 MKGWMTALLGALIGTAWA.....LCSKRTDCHRALHSHDEL 182

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV-xlp  
-O/-cgn2.1/USPTO.spool/US10082502/runat.24032003.135058.6538/app.query.fasta.1.1308  
-DB-N Geneseq 101002 -QFMT-fastap -SUFTX-rng -MINMATCH=0.1 -ICPCL=0  
-LOOEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-BLOSUM62 -TRANS-human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR-SCORE-pct -THR-MAX=100 -THR-MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NOR-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10082502.ecgn.1.1.741-@runat.24032003.135058.6538 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOC -DEV TIMEOUT=120  
-NARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq.101002.\*  
1: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	960	100.0	546	AA238326	Human transmembran
2	960	100.0	806	AA206969	Human secretory pe
3	960	100.0	806	AA208293	Human zsl199 gene e
4	960	100.0	814	AA238327	Human transmembran
5	960	100.0	814	AA252765	CDNA encoding tran
6	960	100.0	814	ABL41995	Nucleotide sequenc
7	960	100.0	814	ABK09772	Human ovarian tumo
8	960	100.0	1180	AA221112	Human secreted pro
9	960	100.0	1210	AA221330	Human CDNA sequenc
10	953	99.3	714	AA206970	Human secreted pro
11	938	97.7	1069	AA206970	Mouse secretory pe
12	938	97.7	1069	AA208294	Mouse ortholog gen
13	914	95.2	657	ABK36007	CDNA sequence #398
14	911	94.9	832	AA297837	Human secreted pro
15	835.5	87.0	592	AA443525	Human secreted exp
16	393	40.9	484	AA125628	Human breast cance
17	393	40.9	558	AA118023	Human breast cance
18	393	40.9	595	AA107882	Human breast cance
19	393	40.9	1160	AA126689	Human breast cance
20	369	38.4	649	AA297883	Human secreted pro
21	369	38.4	649	AA206968	Human secretory pe
22	369	38.4	649	AA208284	Human zsl199 gene.
23	369	38.4	718	AA210655	CDNA encoding a hu
24	366	38.1	792	AA297836	Human secreted pro
25	365	38.0	415	AA206971	zsl199 expressed se
26	365	38.0	415	AA208285	Expressed Sequence
27	355	37.0	1085	AA221330	WO9901020 Seg ID 3
28	316	32.9	2532	ABA07341	Human pancreatic c
29	316	32.9	2532	AA232768	Human genomic DNA
30	316	32.9	2724	ABA07340	Human pancreatic c
31	316	32.9	2724	AA232767	Human genomic DNA
32	266	27.7	564	ABK09691	Human ovarian tumo
33	212	22.1	816	ABL21471	Drosophila melanog
34	184	19.2	352	ABA09488	Human secreted pro
35	154	16.0	786	AA237767	Arabidopsis thalia
36	147	15.3	2750	ABL16485	Drosophila melanog
37	147	15.3	2750	ABL16484	Drosophila melanog
38	147	15.3	3274	ABL16482	Drosophila melanog
39	141	14.7	3110	ABL21470	Drosophila melanog
40	131.5	13.7	744	AA249678	Human myocardiun s
41	131.5	13.7	746	AA249678	Human myocardiun s
42	131.5	13.7	747	AA249678	Human myocardiun s
43	131.5	13.7	747	AA249678	Human myocardiun s
44	131.5	13.7	747	AA249678	Human myocardiun s
45	131.5	13.7	747	AA249678	Human myocardiun s

## ALIGNMENTS

RESULT 1  
AA238326  
AA238326 standard; CDNA; 546 BP.

AA238326:  
09-FEB-2000 (first entry)  
Human transmembrane protein CDNA clone HP10390 coding sequence.

HP10390: transmembrane domain; stomach cancer cell; antibody;  
assay reagent; diagnostic marker; primer; probe; antisense; gene therapy;  
agonist; antagonist; ligand; therapeutic; ds.

OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..540

```

FT      /tag-a
FT      /product= "Human transmembrane protein HP10390"
FT      /note= "No stop codon given in the specification"
XX
XX
XX      MO9955862-A2.
XX
XX      04-NOV-1999.
XX
XX      27-APR-1999;      99WO-JP02226.
XX
XX      28-APR-1998;      98JP-0119395.
XX
XX      (SAGA ) SAGAMI CHEM RES CENT.
XX      (PROT-) PROTEGENE INC.
XX
XX      Kato S, Kimura T;
XX
XX      WPI: 2000-023358/02.
XX      P-PSDB: AAY52391.
XX
XX      Human proteins with transmembrane domains, involved in control of cell
XX      proliferation and differentiation, useful for treating e.g. cancer or
XX      inflammation
XX
XX      Claim 3; Page 88; 114pp: English.
XX
XX
XX      This sequence represents the coding sequence of human cDNA clone
XX      HP10390 which encodes a 20 kD protein with one putative transmembrane
XX      domain in the N-terminus. The cDNA was isolated from a human stomach
XX      cancer cell line cDNA library. The protein has no homology with any
XX      known protein. The protein may be used to raise specific antibodies, as
XX      assay reagents, as diagnostic tissue markers, for the isolation of
XX      cognate receptors, ligands and binding proteins, and as biologically
XX      active agents. Nucleotides encoding the protein may be used as primers
XX      and probes or antisense molecules, and in gene therapy. Cells transformed
XX      with these nucleotides may be used to screen for agonists and antagonists
XX      which are potentially useful therapeutically.
XX
XX      Sequence 546 BP; 139 A; 126 C; 165 G; 116 T; 0 other;
XX
XX
XX      Alignment Scores:
XX      Pred. NO.:      1.48e-103      Length:      546
XX      Score:      960.00      Matches:      182
XX      Percent Similarity:      100.00%      Conservative:      0
XX      Best Local Similarity:      100.00%      Mismatches:      0
XX      Query Match:      100.00%      Indels:      0
XX      DB:      21      Gaps:      0
XX
XX      US-10-082-502-17 (1-182) x AAZ38326 (1-546)
XX
XX      Oy      1      MetLvgLyTRpGLyTRPLeuAlaLeuLeuLeuGlyAlaLeuLeuGlyThrAlaTPala 20
XX      Db      1      ATGAAAGGCTGGGGTGGCTGGCCCTCGCTTCGGGGGCCCTGCTGGGAACCGCCTGGGCT 60
XX
XX      Oy      21      ATGATGSerGlnAspLeuHisCysGlyAlaCysATGAlaLeuValAspGluLeuGluTTP 40
XX      Db      61      CGAGAGAGCAAGATCTCCACTGTGGAGCATCGAGGGCTCTGGTGGATGAATGATGG 120
XX
XX      Oy      41      GluileAlaGlnValaAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60
XX      Db      121      GAAATGTGCCAGCTGAGACCCCAAGAAGACCATTCACATGGAGATCTTCCCGATCAATCCA 180
XX
XX      Oy      61      AspGlySerGlnSerValaIgluValProTyraIaArgSerGlnAlaHisLeuThrGlu 80
XX      Db      181      GATGCAACCACTCAGTGTGGTGGAGGTGCTTATGCCCCGCTCAGAGGCCACCTCACAGAG 240
XX
XX      Oy      81      LeuLeuGluGluIleCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 100
XX      Db      241      CTGCTGGAGGAGATATGTGACCGGATGAAGGAGTATGGGGGAACAGATTGATCTTGCACC 300
XX
XX      Oy      101      HisArgLysAsnTyValaArgValaGlyAaGaaGluGluSerSerGluLeuAspLeu 120
XX      Db      301      CATGCCAAGAACTACGTACGTGTAGTGGGCCGAATGGAGAAATCCAGTAACGTGACCTTA 360

```

Db	361	CAGAGCATCCGAAACGACGCTCAGATATTAGGGGACCCCTCAAGTTTGGGTGAGAGCAT	420
Qy	121	GlnGlyIleArgIleAspSerArgIleSerGlyThrIleLysPheAlaCysGluSerIle	140
Qy	121		
Db	421	GTGAGAGAAATACGAGCGATGAACTCATTTGATCTTTCCCGAGAGGCTGCATATCTTAA	480
Qy	141	ValGluGluTrpGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys	160
Qy	161	AspLysLeuCysSerLysAspGluThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp	180
Qy	181		
Db	481	GACAAACTTTCGACAGTAAAGCAACAGATCTTTGTGACCATGCGCTCCACATATATGCATGAT	540
Db	541	GAGCTA 546	
RESULT 2			
AAx06969	AAx06969 standard; cDNA; 806 BP.		
AAx06969;			
AC			
DT	10-MAY-1999 (first entry)		
XX			
XX	Human secretory peptide-9 (Zs1g9) variant cDNA.		
DE			
XX			
KW	Secretory peptide-9; Zs1g9; human; tumour marker; cancer; therapy;		
KW	diagnosis; growth enhancer; variant; ss.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	CDS	104..652	
FT		/*tag- a	
FT	sig_peptide	104..163	
FT		/*tag- b	
FT	mat_peptide	164..649	
FT		/*tag- c	
XX			
XX	W09901554-A1.		
XX			
PD	14-JAN-1999.		
XX			
PF	02-JUL-1998;	98MO-US13859.	
XX			
PR	17-JUN-1998;	98US-0099005.	
PR	03-JUL-1997;	97US-0051704.	
PR	03-JUL-1997;	97US-0888088.	
PR	19-MAY-1998;	98US-0081338.	
PR	19-MAY-1998;	98US-0085983.	
PR	17-JUN-1998;	98US-0089899.	
XX			
PA	(ZYMO ) ZYMOGENETICS INC.		
PI	Jaspers SR, Jelinek LJ, Sheppard PO, Whitmore TE;		
XX			
XX	WPI: 1999-106055/09.		
DR	P-PSDB; AAM88474.		
DR			
XX			
FT	New mammalian secretory peptide-9 (Zs1g9) - used as a growth		
PT	enhancer for placenta, liver and heart, and as an indicator of		
PT	cancer		
XX			
PS	Claim 2; Page 73-74; 85pp; English.		
XX			
CC	This cDNA clone encodes human secretory peptide-9, or Zs1g9,		
CC	variant (see AAM88474). Zs1g9 (see also AAM88469) is overexpressed in		
CC	human brain, liver, lung, oesophageal, stomach, colon, rectal,		
CC	thyroid and lymphoma tumors. Thus, Zs1g can be used as an		
CC	indicator for cancer. Zs1g9 cDNA was discovered in a placenta		
CC	clone from a full-term pregnancy cDNA library which contained an		
CC	expressed sequence tag (see AAX06971). The invention provides		

CC polynucleotides (see AAX06968-70) encoding Zs199 polypeptides (see  
 CC AAM88469-77) including mature polypeptides, other processed forms,  
 CC variants and mouse orthologues. The Zs199 gene, or probes derived  
 CC from it, can be used to determine if Zs199 is present on chromosome  
 CC 12, and if a mutation has occurred. Antibodies raised against  
 CC Zs199 can be used as diagnostic agents to determine the presence of  
 CC Zs199, and thus the presence of cancer. They can also be labelled  
 CC with radiolabels or fused with toxins and used to treat tumours  
 CC which overexpress Zs199. Antisense nucleotides derived from Zs199  
 CC cDNA can also be used to inhibit the growth of tumour cells. Zs199  
 CC proteins can be used to enhance the growth or development of the  
 CC placenta, heart or liver.

XX  
 SQ Sequence 806 BP; 229 A; 183 C; 232 G; 160 T; 2 other:

# Alignment Scores:

Pred. No.:	2,49e-103	Length:	806
Score:	960.00	Matches:	182
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	20	Gaps:	0

US-10-082-502-17 (1-182) x AAX06969 (1-806)

QY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTPala 20  
 DB 104 ATGAAAGCGTGGGCTGGCGCTGCTCTGCGGCGCTGCGGAAACCGCTGGGCT 163  
 QY 21 ArgArgSerGlnAspLeuHisGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 DB 164 CGAGAGAGCCAGATCTCCCTGAGCATGCAAGGCTGTGGTGAAGTGAAGATGG 223  
 QY 41 GluIleAlaGlnValAspProLysTrpThrIleGlnMetGlySerPheArgIleAspPro 60  
 DB 224 GAATATGCGCCAGGTGGACCCCAAGAACCATTCAGATGGATCTTCGAGCATTCAC 283  
 QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
 DB 284 GATGGCAGCCAGTCAGTGGGAGGTGCTTATGCTCCGCTCAAGGCCACCTCACAGAG 343  
 QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
 DB 344 CTCCTGAGAGAGATATGTGACCGGATGAAGAGATATGGGACAGATTCCTTCACCC 403  
 QY 101 HisArgLysAsnTyrValArgValValGlyArgGlnGlyLysSerGluLeuAspLeu 120  
 DB 404 CATTGCGAAGAACTACGTAGTGGGCGGGAATGGAATCCAGTGAACGACCTTA 463  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 464 CAAAGCATCCGAATCGACCTGATATTACGGCACCCTCAAGCTTCCGTGAGACCAT 523  
 QY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
 DB 524 GTGAGAGATAGACAGATGATGATGATTTTCCCGAGAGCTGACATGTTAA 583  
 QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180  
 DB 584 GACAAACTTGGAGTAAAGCAACAGATCTTGGACCATGCCCTGCATATCGCATGAT 643  
 QY 181 GluLeu 182  
 DB 644 GACCTA 649

## RESULT 3

AAZ08293  
 ID AAZ08293 standard; DNA; 806 BP.

AC AAZ08293;

DT 07-FEB-2000 (first entry)

XX

DE Human Zs199 gene encoding secretory protein variant-4.

XX  
 KW Secretory protein-9; Human Zs199; chromosome 12q15 region; variant;  
 KW overexpression; antagonist; antibody; antisense nucleotide; tumour;  
 KW treatment; receptor; radio-label; fusion; polypeptide toxin; technique;  
 KW down-regulation; probe; diagnostic; therapeutic; cancer; brain; liver;  
 KW detection; stomach; lymphoma; alternative splicing; allelic variation;  
 KW silent mutation; ds.

XX  
 OS Homo sapiens.  
 OS Synthetic.

XX  
 FH Key Location/Qualifiers  
 FT CDS 104..652

FT /tag= "Zs199 secretory protein variant-4"  
 FT /note= "Overexpressed in tumours"

FT sig\_peptide 104..163  
 FT /tag= b  
 FT mat\_peptide 164..649  
 FT /tag= c  
 FT /label= Mature\_Zs199\_protein\_variant-4

XX  
 PN W09960405-A1.

XX  
 PD 25-NOV-1999.

XX  
 PF 19-MAY-1999; 99WO-US11107.

XX  
 PR 19-MAY-1998; 98US-0081183.

XX  
 PA (ZYMO ) ZYMOGENETICS INC.

XX  
 PI Moore BE, Taft DW;

XX  
 DR WPI. 2000-039447/03.

XX  
 DR P-PSDB; AAY15135.

XX  
 PT Detecting tumors using antibodies, antagonists and antisense  
 PT nucleotides to secretory protein-9 (Zs199)

XX  
 PS Disclosure: Page 33-35; 45pp: English.

XX  
 CC The present DNA sequence is a gene encoding the variant of the secretory  
 CC protein-9, Zs199 that arises due to alternative splicing, allelic  
 CC variation or silent mutations that result in amino acid changes. This  
 CC sequence is mapped to the human chromosome 12q15 region. It is  
 CC overexpressed in tumours. Antagonists, antibodies and antisense  
 CC nucleotides to Zs199 are useful for detecting and treating tumours. The  
 CC antagonist may be an antibody or receptor to Zs199 and it may be radio-  
 CC labelled or fused to a polypeptide toxin. It can be used for down  
 CC regulating the overexpression of Zs199. The gene sequence can be used as  
 CC nucleic acid probes to detect RNA encoding Zs199. The Zs199 sequence  
 CC facilitates improved diagnostic and therapeutic techniques for detecting  
 CC and treating cancers, especially of the brain, liver, stomach, lymphoma,  
 CC etc., at an early stage.

XX  
 SQ Sequence 806 BP; 229 A; 183 C; 233 G; 160 T; 1 other:

## Alignment Scores:

Pred. No.:	2,49e-103	Length:	806
Score:	960.00	Matches:	182
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-10-082-502-17 (1-182) x AAZ08293 (1-806)

QY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTPala 20  
 DB 104 ATGAAAGCGTGGGCTGGCGCTGCTCTGCGGCGCTGCGGAAACCGCTGGGCT 163

```

OY 21 ATGAGSerglnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40
DB 164 CGGAGAGCCAGATCTCCACTGTGAGCATGACAGGCGCTGTGTGATGAACTAGAAATGG 223
OY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60
DB 224 GAAATTTGCCAGGTGACCCCAAGAACATTCAGATGGATCTTTCCGGATCAATCCA 283
OY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
DB 284 GATGACAGCCAGTCACTGTGTGGAGGTGCTTATGCCCGCTCAGAGGCCACCTCACAGAG 343
OY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100
DB 344 CTGCTGGAGAGATATGTACCGGATGAAGAGTATGGGAAACAGATTGATCTTCACAC 403
OY 101 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerSerGluLeuAspLeu 120
DB 404 CATCCCAAGAACTAGCTAGCTAGTGTAGTGGCCGGAATGGGAATCCAGTCAACTGGACCTA 463
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
DB 464 CAAGGCATCCGAAATGACATGATATAGCGGACCCCTCAAGTTGGCGTGAGAGCATT 523
OY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160
DB 524 GTGGAGGAATACGAGATGAATCAATGAAATCTTTTCCGAGAGGCTCACAATGTGTA 583
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180
DB 584 GACAAACTTTGAGTAAAGCGAACACATCTTTGTGACATGCCCTGCACATATCGCATGAT 643
OY 181 GluLeu 182
DB 644 GAGCTA 649

RESULT 4
AAZ38327 standard; cDNA: 814 BP.
AC AAZ38327;
DT 09-FEB-2000 (first entry)
DE Human transmembrane protein cDNA clone HP10390.
XX
KW HP10390; transmembrane domain; stomach cancer cell; antibody;
KW assay reagent; diagnostic marker; primer; probe; antisense; gene therapy;
KW agonist; antagonist; ligand; therapeutic; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 145..693
FT /tag="a
FT /product="Human transmembrane protein HP10390"

WO9955862-A2.
XX 04-NOV-1999.
XX 27-APR-1999; 99WO-JP02226.
XX
XX 28-APR-1998; 98JP-0119395.
XX (SAGA ) SAGAMI CHEM RES CENT.
XX (PROT-) PROTEGENE INC.
XX Kato S, Kimura T;
XX
XX WPI: 2000-023358/02.
XX P-PSDB; AAY52391.
XX

```

```

PT Human proteins with transmembrane domains, involved in control of cell
PT proliferation and differentiation, useful for treating e.g. cancer or
PT inflammation
XX
XX Claim 4; Page 106-107; 114pp; English.
XX
CC This sequence represents the human cDNA clone HP10390
CC which encodes a 20 kD protein with one putative transmembrane
CC domain in the N-terminus. The cDNA was isolated from a human stomach
CC cancer cell line cDNA library. The protein has no homology with any
CC known protein. The protein may be used to raise specific antibodies, as
CC assay reagents, as diagnostic tissue markers, for the isolation of
CC cognate receptors, ligands and binding proteins, and as biologically
CC active agents. Nucleotides encoding the protein may be used as primers
CC and probes or antisense molecules, and in gene therapy. Cells transformed
CC with these nucleotides may be used to screen for agonists and antagonists
CC which are potentially useful therapeutically.
SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Alignment Scores:
Pred. No.: 2,53e-103 Length: 814
Score: 960.00 Matches: 182
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 21

US-10-082-502-17 (1-182) x AAZ38327 (1-814)
OY 1 MetLysGlyTyrPglYThrPleuAlaLeuLeuGluAlaLeuGlyThrAlaTrpAla 20
DB 145 ATGAAGGCTGGGGTGGCTGGCCCTCTCTCTGGGGGCGCTGCGAACCCTGGGCT 204
OY 21 ArgArgSerglnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40
DB 205 CGGAGAGCCAGGATCTCCACTGTGAGCATGACAGGCTCTGTGTGATGAATAGAAATGG 264
OY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60
DB 265 GAAATTTGCCAGGTGACCCCAAGAACATTCAGATGGATCTTTCCGGATCAATCCA 324
OY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
DB 325 GATGACAGCCAGTCACTGTGTGGAGGTGCTTATGCCCGCTCAGAGGCCACCTCACAGAG 384
OY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100
DB 385 CTGCTGGAGAGATATGTGACCGGATGAAGAGTATGGGAAACAGATTGATCTTCACAC 444
OY 101 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerSerGluLeuAspLeu 120
DB 445 CATCCCAAGAACTAGCTAGCTAGTGTAGTGGCCGGAATGGGAATCCAGTCAACTGGACCTA 504
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
DB 505 CAAGGCATCCGAAATGACATGATTTGCGGCGCCCTCAAGTTTGGCTGTGAGAGCATT 564
OY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160
DB 565 GTGGAGGAATACGAGATGAATCAATGAAATCTTTTCCGAGAGGCTGACAAATGTGTA 624
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180
DB 625 GACAAACTTTGAGTAAAGCGAACACATCTTTGTGACATGCCCTGCACATATCGCATGAT 684
OY 181 GluLeu 182
DB 685 GAGCTA 690

RESULT 5
ABK52765
ID ABK52765 standard; cDNA: 814 BP.

```

XX	ABK52765;			
XX	15-AUG-2002 (first entry)			
XX	CDNA encoding transmembrane protein 4, a cancer-linked protein.			
XX	Expressed sequence tag; EST; human; cancer; anti-neoplastic; cytoxic T lymphocyte; chemotherapy; cytosstatic; chromosome 12; transmembrane protein 4; gene; ss.			
XX	Homo sapiens.			
XX	Key	location/Qualifiers		
FT	CDS	145..693		
FT		/*tag= a		
FT		/product= "transmembrane protein 4"		
PN	W0200231198-A2.			
XX	18-APR-2002.			
XX	11-OCT-2001; 2001WO-US31607.			
XX	11-OCT-2000; 2000US-239294P.			
PR	11-OCT-2000; 2000US-239297P.			
PR	11-OCT-2000; 2000US-239605P.			
PR	12-OCT-2000; 2000US-239802P.			
PR	12-OCT-2000; 2000US-239805P.			
PR	12-OCT-2000; 2000US-239806P.			
PR	16-OCT-2000; 2000US-240622P.			
PR	19-OCT-2000; 2000US-241682P.			
PR	19-OCT-2000; 2000US-241723P.			
PR	31-OCT-2000; 2000US-244932P.			
PA	(AVAL-) AVALON PHARM.			
PI	Young PE, Horrigan S, Weaver Z, Endress GA;			
XX	WPI; 2002-4632771/49.			
DR	P-PSDB; AAU97063.			
PT	Identifying modulators of a cancer-related gene to screen agents for preventing or treating cancer comprises detecting a difference in the expression of cancer-linked genes in the presence or absence of test compounds			
XX	Claim 1; Page 43; 66pp; English.			
PS	The invention relates to modulators of a cancer-related genes. Also described are: (1) processes for identifying an anti-neoplastic agents comprising contacting a cell exhibiting neoplastic activity with a compound first identified as a cancer related gene modulator, and detecting a decrease in the neoplastic activity. (2) a process for determining the cancerous state of a cell by determining an increase in the level of expression of at least one gene, where an elevated expression relative to a known non-cancerous cell indicates a cancerous state or potentially cancerous state. The anti-neoplastic agent is useful for treating cancer or for protecting an animal against cancer. The immunogenic composition is also useful for treating cancer in an animal, where the composition elicits the production of cytotoxic T lymphocytes specific for the immunogenic composition. Preferably, the animal is a human. The cancer-linked genes and polypeptides are also useful as targets for cancer therapy or chemotherapy. The present sequence represents a cancer-linked gene located on chromosome 12, which encodes transmembrane protein 4.			
XX	Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;			
XX	Alignment Scores:			
XX	Pred. No.:	2,53e-103	Length:	814
XX	Score:	960.00	Matches:	182
XX	Percent Similarity:	100.00%	Conservative:	0

[illegible]



QY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20  
 DB 145 ATGAAAGGCTGGGTGGTGGCTGCTTCTGGGGGGCCCTGCTGGAAACCGCTGGCT 204  
 QY 21 ArgArgSerGlnAspLeuHicysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 DB 205 CGAGAGACCAAGGATCTCAGCTGTCGACATCCAGGCGCTTGGTGGATGAACCTGAATGG 264  
 QY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 265 GAAATGGCCAGGTGGACCCCAAGAAACATTCAGATGGGATCTTCCGATCATATCA 324  
 QY 61 AspGlySerGlnSerValAlaGluValProGlyAlaAspSerGluAlaHisLeuThrGlu 80  
 DB 325 GATGGACACCCAGTCAGTGGTGGAGGTGCTTATGCCCTCAGAGGCCACCTCACAAG 384  
 QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThr 100  
 DB 385 CTGCTGGAGAGATATGTGACCGGATGAAGAGATGGGAAACAGATTGATCTTCCACC 444  
 QY 101 HisArgLysAsnTrpValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 120  
 DB 445 CATGGCAAGAACTACGACGTGTAGTGGCCCGAATGGAGAAATCCAGTGAACCTGACCTA 504  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 505 CAAGGATCCGAACTCGACTCGATATTATGAGCGCACCTCAAGTTGGTGTGAGACGATT 564  
 QY 141 ValGluGluTrpGluAspGluLeuIleGluPhePheSerArgLysAlaAspAsnValLys 160  
 DB 565 GTGGAGGAATACGAGATGAACTGATTCGANTCTTTCCGAGAGCGCTGACATGTTTAA 624  
 QY 161 AspLysLeuGlySerLysValArgThrAspLeuLysAspHisAlaLeuHisIleSerHisAsp 180  
 DB 625 GACAACTTGCAGTAAAGCAACGATCTTGTGACCATGCTCCCTGACATATGTCATGAT 684  
 QY 181 GluLeu 182  
 DB 685 GAGCTA 690  
 RESULT 8  
 AAX22112  
 ID AAX22112 standard; DNA; 1180 BP.  
 AC AAX22112;  
 XX  
 DT 18-MAY-1999. (first entry)  
 XX  
 DE Human secreted protein gene 2 clone H2MB56.  
 XX  
 KW Human; secreted protein; gene therapy; protein therapy; tissue; cancer;  
 KW tumour; neurodegenerative disorder; leukemia; autoimmune disease; AIDS;  
 KW developmental abnormality; foetal deficiency; Alzheimer's disease;  
 KW cognitive disorder; schizophrenia; immunological disorder; mood disorder;  
 KW immune deficiency disease; respiratory disorder; arthritis; skeletal;  
 KW haematopoietic disorder; neural; osteoporosis; metabolic disorders;  
 KW cardiovascular; endocrine; gastrointestinal; asthma; diagnosis; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9901020-A2.  
 XX  
 PD 14-JAN-1999.  
 XX  
 PF 30-JUN-1998; 98MO-US13608.  
 XX  
 PR 12-SEP-1997; 97US-0058663.  
 PR 01-JUL-1997; 97US-0051381.  
 PR 01-JUL-1997; 97US-0051480.  
 PR 12-SEP-1997; 97US-0058598.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.

PI Carter KC, Endress GA, Feng P, Rosen CA, Ruben SM;  
 XX WPI: 1999-105683/09.  
 DR P-PSDB: AAY01136, AAY01162, AAY01163.  
 PT  
 PT New isolated human genes and the secreted polypeptides they encode -  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, immune deficiency diseases or blood  
 PS disorders  
 PS Claim 4; Page 117; 179pp; English.  
 XX  
 XX The invention relates to nucleic acid sequences (AAX22111 to AAX22124)  
 CC encoding human secreted proteins (AAY01135 to AAY01158). The secreted  
 CC protein gene sequences are deposited with the ATCC under deposit number  
 CC ATCC 209118. Host cells comprising recombinant vectors containing the  
 CC nucleic acid sequences are used for the recombinant production of the  
 CC secreted proteins. The polynucleotide and amino acid sequences of the  
 CC for are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Pathological conditions can  
 CC be also diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new  
 CC polynucleotides. Specific uses are described for each of the  
 CC CC In, and include developing products for the diagnosis or treatment of  
 CC cancer, tumours, developmental abnormalities and foetal deficiencies,  
 CC autoimmune diseases, lymphomas, Alzheimer's and cognitive disorders,  
 CC schizophrenia, immunological disorders, immune deficiency diseases  
 CC (AIDS), mood disorders, respiratory disorders, arthritis, asthma,  
 CC haematopoietic disorders, neural disorders, skeletal disorders,  
 CC osteoporosis or gastrointestinal disorders, cardiovascular disorders, endocrine  
 CC disorders or gastrointestinal disorders. The polypeptides are also useful  
 CC for identifying their binding partners. The present sequence represents a  
 CC gene encoding a human secreted protein (see descriptor line for gene  
 CC number and clone identification).  
 XX  
 SQ Sequence 1180 BP; 258 A; 304 C; 363 G; 250 T; 5 other:  
 Alignment Scores:  
 Pred. No.: 4, 16e-103 Length: 1180  
 Score: 960.00 Matches: 182  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0  
 US-10-082-502-17 (1-182) x AAX22112 (1-1180)  
 QY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20  
 DB 513 ATGAAAGGCTGGGTGGTGGCTGCTTCTGGGGGGCCCTGCTGGAAACCGCTGGCT 572  
 QY 21 ArgArgSerGlnAspLeuHicysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 DB 573 CGAGAGACCAAGGATCTCAGCTGTCGACATCCAGGCGCTTGGTGGATGAACCTGAATGG 632  
 QY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 633 GAAATGGCCAGGTGGACCCCAAGAAACATTCAGATGGGATCTTCCGATCATATCA 692  
 QY 61 AspGlySerGlnSerValAlaGluValProGlyAlaAspSerGluAlaHisLeuThrGlu 80  
 DB 693 GATGGACACCCAGTCAGTGGTGGAGGTGCTTATGCCCTCAGAGGCCACCTCACAAG 752  
 QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThr 100  
 DB 753 CTGCTGGAGAGATATGTGACCGGATGAAGAGATGGGAAACAGATTGATCTTCCACC 812  
 QY 101 HisArgLysAsnTrpValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 120  
 DB 813 CATGGCAAGAACTACGACGTGTAGTGGCCCGAATGGAGAAATCCAGTGAACCTGACCTA 872  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140

```
DB 873 CAAGGCATCCGATGACATGATATAGCGGACCCCTCAACTTTGCGTGAGAGCATT 932
OY 141 ValGluGluTYrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnVallys 160
DB 993 GTGGAGGAAATACGAGATGAACTCATGTATCTTTCCCGAGAGCGTCACAAATCTTAA 992
OY 161 AspIlyLeuCYSerSerLysArgThrAspLeuCYAspHisAlaLeuHisIleSerHisasp 180
DB 993 GACAAACTTTCAGTAGACGAGACAGATCTTGTGTACATAGCCCTGCACATATCGCATGAT 1052
OY 181 GluLeu 182
DB 1053 GAGCTA 1058

RESULT 9
AAS21330
ID AAS21330 standard; cDNA; 1210 BP.
XX
AC AAS21330;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human cDNA sequence encoding for PRO4426 polypeptide.
XX
KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW breast; prostate; cervical; tumour necrosis factor- $\alpha$ ; TNF- $\alpha$ ;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200140466-A2.
XX
PD 07-JUN-2001.
XX
PE 01-DEC-2000; 2000WO-US32678.
XX
PF 01-DEC-1999; 99WO-US28301.
XX
PR 01-DEC-1999; 99WO-US28634.
XX
PR 02-DEC-1999; 99WO-US28551.
XX
PR 02-DEC-1999; 99WO-US28564.
XX
PR 09-DEC-1999; 99US-0170262.
XX
PR 16-DEC-1999; 99WO-US30095.
XX
PR 20-DEC-1999; 99WO-US30911.
XX
PR 20-DEC-1999; 99WO-US30999.
XX
PR 30-DEC-1999; 99WO-US31243.
XX
PR 06-JAN-2000; 2000WO-US00277.
XX
PR 06-JAN-2000; 2000WO-US00376.
XX
PR 11-FEB-2000; 2000WO-US03565.
XX
PR 18-FEB-2000; 2000WO-US04341.
XX
PR 18-FEB-2000; 2000WO-US04342.
XX
PR 22-FEB-2000; 2000WO-US04914.
XX
PR 24-FEB-2000; 2000WO-US04914.
XX
PR 24-FEB-2000; 2000WO-US05004.
XX
PR 01-MAR-2000; 2000WO-US05601.
XX
PR 20-MAR-2000; 2000WO-US07377.
XX
PR 21-MAR-2000; 2000WO-US07532.
XX
PR 30-MAR-2000; 2000WO-US08439.
XX
PR 17-MAY-2000; 2000WO-US13705.
XX
PR 22-MAY-2000; 2000WO-US14042.
XX
PR 30-MAY-2000; 2000WO-US14941.
XX
PR 02-JUN-2000; 2000WO-US15264.
XX
PR 10-NOV-2000; 2000WO-US30873.
XX
PA (GETH ) GENENTECH INC.
XX
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,
PI Gerlisen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S,
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2001-408281/43.
```

```
DR P-PSDB; AAU12258.
XX
XX Isolated, secretory and transmembrane PRO polypeptide used to detect
PT other PRO polypeptides, link bioactive molecules to cells expressing
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT lung, breast, prostate, cervical
XX
XX Claim 3; Fig 173; 813pp; English.
PS
XX
XX AAS21244-AAS21518 encode for novel human secretory and transmembrane
CC PRO polypeptides. The PRO polypeptides are useful to detect other
CC PRO polypeptides, to link bioactive molecules to cells expressing
CC PRO polypeptides, to modulate biological activities of cells expressing
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC polypeptide expression in a cell sample to that in a control sample.
CC Some of the 275 sequences are also useful to stimulate the release of
CC tumour necrosis factor- $\alpha$  (TNF- $\alpha$ ) from human blood, the
CC proliferation or differentiation of chondrocytes, the proliferation or
CC gene expression in pericyte cells, the release of proteoglycans from
CC cartilage, the proliferation of inner ear utricular supporting cells or
CC of T-lymphocytes, the release of a cytokine from peripheral blood
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC to factor VIIA. The PRO polypeptides can be used in assays to identify
CC molecules involved in binding interactions. The polynucleotides encoding
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy.
XX
XX Sequence 1210 BP; 261 A; 316 C; 379 G; 254 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 4,31e-103 Length: 1210
Score: 960.00 Matches: 182
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-10-082-502-17 (1-182) x AAS21330 (1-1210)
OY 1 MetLysGlyTTPbGlyTTPLeuAlaLeuLeuLeuGlyAlaLeuLeuGlyThrAlaTrrAla 20
DB 543 ATGAAGGCTGGGGTGGCTGGCCCTGCTTGGGGGGCCCTGGGAACCGCCGGGCT 602
OY 21 ArgArgSerGlnAspLeuHisCYSerGlyAlaCYAspArgAlaLeuAlaAspGluTrp 40
DB 603 CGGAGAGCCAGGATCTCCACTGCTGGACATCGACGGCTCTGGTGATGAACCTAGAAATGG 662
OY 41 GluIleAlaGluValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60
DB 663 GAAATTTGCCAGGTCGACCCCAAGAAAGACCATTCAGATGGGATCTTCCGGATCAATCCA 722
OY 61 AspGlySerGlnSerValValGluValProGlyTrpAlaArgSerGluAlaHisLeuThrGlu 80
DB 723 GATGGCACCCAGTCAGTGTGGAGGTGCTTATGCGCCCTCAGAGGCCACCTCACAGAG 782
OY 81 LeuLeuGluGluIleCYAspArgMetLysGlyIlyrGlyGluGlnIleAspProSerThr 100
DB 783 CTGCTGGAGGAGATRTGTGACCGGATGAGAGATRTGGGAAACAGATGATCTTCCACC 842
OY 101 HisArgLysAsnTYrValArgValValGluValArgAsnGlyIleSerSerGluLeuAspLeu 120
DB 843 CATGCAAGAACTACGTAGTGTAGTGGCCGGAATGAGAAATCAGTAAGTGGACCTA 902
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCYGluSerIle 140
DB 903 CAAGGCATCCGATGACATGATATAGCGGACCCCTCAAGTTTGGCTGTGAGAGCATT 962
OY 141 ValGluGluTYrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnVallys 160
DB 963 GTGGAGGAAATACGAGATGAACTCATGTATCTTTCCCGAGAGCGTCACAAATCTTAA 1022
```



```

QY 161 Asplysleucyserylserlysarqthhraspleucysasphialaleuhtisleserhisasp 180
DB 1023 GACAAACTTTGACATAGCCAGACATCTTGTGACCATGCCCCGACATATCGCATGAT 1082
QY 181 Glutreu 182
DB 1083 GAGCTA 1088

RESULT 10
AAK97884 standard; cDNA; 714 BP.
AC AAK97884;
AT 23-SEP-1999 (first entry)
DE Human secreted protein encoding cDNA #72.
XX
KM Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;
KW diagnostic; gene therapy; chromosome mapping; secretion vector; ss.
XX
OS Homo sapiens.
XX
PN MO9925825-A2.
XX
PD 27-MAY-1999.
XX
PF 13-NOV-1998; 98WO-IB01862.
XX
PR 04-SEP-1998; 98US-0099273.
PR 13-NOV-1997; 97US-0066677.
PR 17-DEC-1997; 97US-0069957.
PR 09-FEB-1998; 98US-0074121.
PR 13-APR-1998; 98US-0081563.
PR 10-AUG-1998; 98US-0096116.
XX
PA (GENE) GENSET.
XX
PI Bougueleret L, Duclert A, Dumas Mline Edwards J;
DR WPI; 1999-347472/29.
DR P-PSDB; AAY36200.
XX
PT Extended cDNAs encoding secreted proteins
XX
PS Claim 1; Page 272; 307pp; English.
XX
CC AAK97813-X97906 represent extended cDNA's which encode novel human
CC secreted proteins (see AAY36129-Y36222) and which have cytosolic,
CC thrombotic and osteopathic activity. The extended cDNAs can be used to
CC express secreted proteins or parts of them or to obtain antibodies
CC capable of binding to the secreted proteins. They may also be used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC Uses also include design of expression vectors and secretion vectors.
XX
SQ Sequence 714 BP; 203 A; 158 C; 202 G; 151 T; 0 other;

Alignment Scores:
Pred. No.: 1,41e-102 Length: 714
Score: 953.00 Matches: 181
Percent Similarity: 99.45% Conservative: 0
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 99.27% Indels: 0
DB: 20 Gaps: 0

US-10-082-502-17 (1-182) x AAK97884 (1-714)
QY 1 MetlysglytyrpglytyrpleualaleuLeuenglyAlaleuenglyThralatrpAla 20
DB 33 ATGAAAGGCTGGGGTGGCTGGCTGCTTCTGGGGGCCCTCTGGAACCCCTGGGCT 92
QY 21 ArgargserglaspleuhtiscysglyalacyarsargalaleuvalaspleuLeuenglyutr 40

```

```

DB 93 CGAGAGAGCCAGATATCCAGCTGTGGAGCATGAGGGCTCTGGTGCATGACATGAATG 152
QY 41 GluIIAlaIaIValaspprolylsytrhileGlmetylserpheargileaspro 60
DB 153 GAAATTTGCCAGGTGGACCCCAAGAACCATTCAGATGGATCTTCCGATCAATCA 212
QY 61 AspglyserglinservValIguValProtyAlaargsergluAlahisLeuthrglu 80
DB 213 GATGGCAGCCAGTCAGTGTGGAGGTGCTTATGCCGCTCAGAGGCCACCTCAGAGAG 272
QY 81 LeuLeuGlugluIuileCysasparqmetlysglutryglygluIuIleaspproserThr 100
DB 273 CTGCTGAGAGATATGTGACCGGATGAAGAGATATGGGGAACAGATTATCTTCCAC 332
QY 101 HisArglyAsnArgValargValValIglyArganglygluIuIleaspproserThr 120
DB 333 CATCGCAGAACTACCTAGCTGTAGTGGCCCGGAAGGAATCCAGTCACTGGACCTA 392
QY 121 GlnGlyIleArgIleasparqmetlysglytyrThrLeuIuIleaspproserThr 140
DB 393 CAAAGCATCGAATGACATCAGATATTAAGCGGACCCCTCAAGTTTCCGTGTGGAGCAT 452
QY 141 ValGluGluTyrGluaspGluLeuIleGluIuIleGluIuIleGluIuIleGluIuIle 160
DB 453 GTGGAGGAATACGAGGATGACATCAATTAATCTTCCGAGAGGCTGACATGTAA 512
QY 161 Asplysleucyserylserlysarqthhraspleucysasphialaleuhtisleserhisasp 180
DB 513 GACAAACTTTGACATAGCCAGACATCTTGTGACCATGCCCCGACATATCGCATGAT 572
QY 181 Glutreu 182
DB 573 GAGCTA 578

RESULT 11
AAK06970 standard; cDNA; 1069 BP.
ID AAK06970;
AC AAK06970;
AT 10-MAY-1999 (first entry)
DE Mouse secretory peptide-9 (Zs1g9) orthologue cDNA.
XX
KW Secretory peptide-9; Zs1g9; orthologue; mouse; tumour marker;
KW cancer; therapy; diagnosis; growth enhancer; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 358..906 /*tag= a
FT sig_peptide 358..417 /*tag= b
FT mat_peptide 418..903 /*tag= c
FT FT
FT FT
XX
PN WO9901554-A1.
XX
PD 14-JAN-1999.
XX
PF 02-JUL-1998; 98WO-US13859.
XX
PR 17-JUN-1998; 98US-0099005.
PR 03-JUL-1997; 97US-0051704.
PR 03-JUL-1997; 97US-0888088.
PR 19-MAY-1998; 98US-0081338.
PR 19-MAY-1998; 98US-0085983.
PR 17-JUN-1998; 98US-0089899.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX

```

PI Jaspers SR, Jellinek LJ, Sheppard PO, Whitmore TE;  
 XX WPI: 1999-106055/09.  
 DR P-PSDB; AAW88476.  
 XX  
 PT New mammalian secretory peptide-9 (Zs19) - used as a growth  
 PT enhancer for placenta, liver and heart, and as an indicator of  
 PT cancer  
 PS Claim 2; Page 75-77; 85pp; English.  
 XX  
 CC This cDNA clone encodes novel mouse secretory peptide-9, or Zs19  
 CC (see AAW88476), an orthologue of novel human Zs19 (see AAW88465).  
 CC Human Zs19 is overexpressed in a number of tumours including  
 CC brain, liver, lung, oesophageal, stomach, colon, rectal, thyroid  
 CC and lymphoma tumors, and thus can be used as an indicator for  
 CC cancer. The invention provides polynucleotides (see AAW6966-70)  
 CC encoding Zs19 polypeptides (see AAW8469-77) including mature  
 CC polypeptides, other processed forms, variants and the mouse  
 CC orthologues. The Zs19 gene, or probes derived from it, can be  
 CC used to determine if Zs19 is present on chromosome 10, and if a  
 CC mutation has occurred. Antibodies raised against Zs19 can be  
 CC used as diagnostic agents to determine the presence of Zs19, and  
 CC thus the presence of cancer. They can also be labelled with  
 CC radioisotopes or fused with toxins and used to treat tumours  
 CC which overexpress Zs19. Antisense nucleotides derived from Zs19  
 CC cDNA can also be used to inhibit the growth of tumour cells. Zs19  
 CC proteins can be used to enhance the growth or development of the  
 CC placenta, heart or liver.  
 XX  
 SO Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other:  
 XX  
 Alignment Scores:  
 Pred. No.: 1,41e-100 Length: 1069  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3  
 Query Match: 97.71% Indels: 0  
 DB: 20 Gaps: 0  
 US-10-082-502-17 (1-182) x AAW6970 (1-1069)  
 OY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20  
 DB 358 ATGAAGGCTGGGTGGGTGGTACCTCTACTTTGGGGTCTCTGGCACTCCGGGCT 417  
 OY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuAlaValAspLeuGlnTrp 40  
 DB 418 CGAAGGAGCGAAGATCTACACTGTGAGCTTGCAGGGCTGTGTGATGATTAAGATGG 477  
 OY 41 GluIleAlaGlnValAspProLysLysThrIleGlnMetClySerPheArgIleAsnPro 60  
 DB 478 GAAATTGCCCGCGTGGAGCCCAAGAACCATTCAGATGGATCCCTCCGATCAATCCA 537  
 OY 61 AspGlySerGlnSerValAlaGluValProTyraAlaArgSerGlnAlaHisLeuThrGlu 80  
 DB 538 GATGGCAGCCACTGATGTTGGAGGTACTTATGCCCTCGAGAGCCACCTCACAGAG 597  
 OY 81 LeuLeuGluGluIleCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThr 100  
 DB 598 TTGCTTGAGGAGGTGTGTGACGATGAAGAGTACGGGGAACAATGACCTTCTTACC 657  
 OY 101 HisArgLysAsnTrpValArgValAlaGlyArgAsnGlyLysSerSerGlnLeuAspLeu 120  
 DB 658 CACCCCAAGAACTACGTACGCTCTGAGCCGAGTGGAGAAATCCAGTAAGACTTAA 717  
 OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 718 CAGGCGATCCGAATGATCATATACAGCGCACCTCAAGTTTGCCTGTGAGAGCATT 777  
 OY 141 ValGluGluTrpGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
 DB 778 GTGGAAGATACGAGGATGAGCTTATCGAATTTCTTCCAGAGAGGCTGACAAAGTTTAA 837

OY 161 AspLysLeuGlySerLysArgThrAspLeuGlyAspHisAlaLeuHisIleSerHisAsp 180  
 DB 838 GACAAACTTTCAGATGAAGCGAGACAGATCTATGTGACCATGCGCCCTGACAGATCTACGAT 897  
 OY 181 GluLeu 182  
 DB 898 GAGCTA 903  
 RESULT 12  
 ID AAZ08294  
 AAZ08294  
 AC AAZ08294; standard: DNA; 1069 BP.  
 AC AAZ08294;  
 DR 07-FEB-2000 (first entry)  
 XX  
 DE Mouse ortholog gene encoding Zs19 secretory protein.  
 XX  
 KW Secretory protein-9; mouse Zs19; ortholog; overexpression; antagonist;  
 KW antibody; antisense nucleotide; tumour; treatment; receptor;  
 KW radio-label; polypeptide toxin; down-regulation; diagnostic;  
 KW therapeutic; probe; cancer; brain; liver; detection; stomach;  
 KW lymphoma; ds.  
 XX  
 OS Mus musculus.  
 XX  
 Key Location/Qualifiers  
 FH 358..906  
 FT CDS  
 FT  
 FT  
 FT sig\_peptide  
 FT /note="Mouse Zs19 secretory protein"  
 FT /note="Overexpressed in tumours"  
 FT /tag= a  
 FT 358..417  
 FT mat\_peptide  
 FT /tag= b  
 FT 418..903  
 FT /tag= c  
 FT /label= Mature\_Zs19\_protein\_variant-4  
 PN W09960405-A1.  
 XX  
 PD 25-NOV-1999.  
 XX  
 PF 19-MAY-1999; 99WO-US11107.  
 XX  
 PR 19-MAY-1998; 98US-0081183.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 PI Moore EE, Taft DW;  
 XX  
 DR WPI: 2000-039447/03.  
 DR P-PSDB; AAY15136.  
 XX  
 PT Detecting tumors using antibodies, antagonists and antisense  
 PT nucleotides to secretory protein-9 (Zs19)  
 PS  
 PS Disclosure; Page 37-38; 45pp; English.  
 XX  
 CC The present DNA sequence is an ortholog encoding the secretory protein-9,  
 CC Zs19 derived from mouse. It is overexpressed in tumours. Antagonists,  
 CC antibodies and antisense nucleotides to Zs19 are useful for detecting  
 CC and treating tumours. The antagonist may be an antibody or receptor to  
 CC Zs19 and it may be radio-labelled or fused to a polypeptide toxin. It  
 CC can be used for down regulating the overexpression of Zs19. The gene  
 CC sequence can be used as nucleic acid probes to detect RNA encoding Zs19.  
 CC The Zs19 sequence facilitates improved diagnostic and therapeutic  
 CC techniques for detecting and treating cancers, especially of the brain,  
 CC liver, stomach, lymphoma etc., at an early stage.  
 SO Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other:  
 Alignment Scores:  
 Pred. No.: 1,41e-100 Length: 1069

Score: 938.00 Matches: 177  
Percent Similarity: 98.358 Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.71% Indels: 0  
DB: 21 Gaps: 0  
US-10-082-502-17 (1-182) x ABK36007 (1-1065)

QY 1 MetlySGlyTTPGLYTPLeuAlaLeuLeuGlyAlaLeuLeuGlyThraAlaTPAla 20  
DB 358 ATGAAAGCTGGGCTGGCTAGCCCTACTTTTGGGGCTCTGCTGGGAAGTGGCTGGCT 417  
QY 21 ArgArgSerGlnAspLeuHscYsgGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
DB 418 CGAAGGAGCCAAAGATCTACACTGTGAGCTGAGGGCTCTGTGTGATGAATTAAGTGG 477  
QY 41 GlnIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 478 GAAATTTGCCGCTGGACCCCAAGAAAGACCATTCAGATGGATCTCTCCGAATCAATCCA 537  
QY 61 AspGlySerGlnSerValAlaGluValProTyraAlaArgSerGlnAlaHscLeuThrGlu 80  
DB 538 GATGGCAGCCAGTCACTGTGGAGTACCTTATGCCGCTCAGAGGCCCACTCAGAGAG 597  
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThr 100  
DB 598 TTGCTTGGAGAGGTGTGTGACCAATGAAGAGACTACGGGCAACAGATTGACCTTCTACC 657  
QY 101 HisArgLysAsnTyraValArgValAlaGlyArgAsnGlyLysSerGluLeuAspLeu 120  
DB 658 CACCGGAAAGTACTACCTAGCGCTGAGCCGGAATGGAAATCCAGTGAATCAATCA 717  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 718 CAGGGCATCCGAATGTACATCAGATATCAGGGCCACCTCAAGTTGCTGTGAGAGATT 777  
QY 141 ValGluGluTrpGluAspGluLeuIleGluPheSerArgLysAlaAspAsnValLys 160  
DB 778 GTGGAAGAATACGAGATGAGCTTATCGAATCTCTCCAGAGAGCTGCACACGTTAAA 837  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHscIleSerHisAsp 180  
DB 838 GACAAACTTTCAGTAAAGCGACAGATCTATGTGACCATGCCCTGCACAGATCTCAGAT 897  
QY 181 GlnLeu 182  
DB 898 GAGCTA 903  
RESULT 13  
ID ABK36007 standard; cDNA; 657 BP.  
AC ABK36007;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE cDNA sequence #398 encoding novel human secreted protein.  
XX  
KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;  
KW Immune deficiency disorder; blood disorder; inflammatory disorder;  
KW Infectious disorder; allergic condition; neurodegenerative disorder;  
KW Liver fibrosis; coagulation disorder; gene therapy; antimicrobial;  
KW tumour; cancer; hepatotropic; immunosuppressive; antineumatic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200177289-A2.  
XX  
PD 18-OCT-2001.  
XX  
PF 29-MAR-2001; 2001WO-US10232.  
XX  
PR 06-APR-2000; 2000US-195605P.

XX (GEMT ) GENETICS INST INC.  
PA Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
PI Merberg D, Treacy M, Agostino MJ, Bowman MR, Spaulding V, Wong CG;  
PI Clark HF, Fectel K, Howes SH, Resnick RJ, Gulukota K, Graham JR;  
XX MPI; 2002-179322/23.  
DR  
XX Six hundred and twenty three polynucleotides derived from a variety of  
XX human tissue sources which encode secreted proteins, useful for  
XX treating immune deficiencies and disorders such as autoimmune disorders  
XX  
XX Claim 1; Page 297; 393pp; English.  
XX  
XX The present invention relates to the isolation of novel cDNA sequences  
XX which encode human secreted proteins. The cDNA sequences have been  
XX derived from a variety of human tissues. The invention also provides  
XX a method for producing proteins from these polynucleotide sequences.  
XX The proteins are useful for identifying compounds that modulate their  
XX activity and production. The sequences of the invention are  
XX useful for treating diseases such as hyperproliferative disorders  
XX (e.g. cancer), immune deficiency disorders (e.g. severe combined  
XX immunodeficiency (SCID)), autoimmune disorders (e.g. multiple  
XX sclerosis), blood disorders (e.g. thrombocytopenia), inflammatory  
XX disorders (e.g. arthritis), infectious disorders (e.g. hepatitis),  
XX allergic conditions (e.g. asthma), neurodegenerative disorders (e.g.  
XX Alzheimer's disease), liver fibrosis, coagulation disorders (e.g.  
XX haemophilia), and tumours. The polynucleotide sequences of the  
XX invention are also useful in gene therapy. ABK3610-ABK36232 represent  
XX the cDNA sequences of the invention that encode for novel human  
XX secreted proteins.  
SQ Sequence 657 BP; 177 A; 149 C; 185 G; 146 T; 0 other;  
Alignment Scores:  
QY 5 GlyTPLeuAlaLeuLeuGlyAlaLeuLeuGlyThraAlaTPAlaArgSerGln 24  
DB 1 GGTGGCTGGCCCTGCTTGTGGGGCCCTGTGGAAACCCCTGGCTCGAGAGACG 60  
QY 25 AspLeuHscYsgGlyAlaCysArgAlaLeuValAspGluLeuGluTrpGlyIleAlaGln 44  
DB 61 GATCTCCACTGTGAGCATCAGGGCTCTGTGATCACTAGAAATGGAAATTTGCCAC 120  
QY 45 ValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnProAspGlySerGln 64  
DB 121 GTGACCCCAAGAAAGACCAAGATGAGATCTTCCGATCAATCAGATTCAGAGCAC 180  
QY 65 SerValValGluValProTyraAlaArgSerGlnAlaHscLeuThrGluLeuGluGlu 84  
DB 181 TCAGTGGTGAAGGTGCTTGTGCCCTGACAGGCCCACTCACAGAGCTGTGGAGAG 240  
QY 85 IleCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThrHisArgLysAsn 104  
DB 241 ATATGTACCCGATGAAGAGATGAGGGAACAGATTCATCTTCCACCCATCCAGAAC 300  
QY 105 TyrValArgValAlaGlyArgAsnGlyLysSerGluLeuAspLeuGlnGlyIleArg 124  
DB 301 TACGTACGTAGTGGCCCGAATGAGAAATCCAGTCACTGACCAAGCAAGCAATCCGA 360  
QY 125 IleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGluTrp 144  
DB 361 ATGACTCAGATATTAGCGGCACTCAAGTTTGGTGTGAGAGATTGGAGAAATAC 420

0Y 145 GUAAPGJLUleuIleGluPheSerArgGluAlaAspAsnValLysAspLysLeuCys 164  
|||||  
Db 421 GAGATGAACATCATTTGAATCTTTCCGAGAGGCTGACAAATGTAAGAACAACCTTTC 480  
0Y 165 SerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAspGluLeu 182  
|||||  
Db 481 AGTAAGCGAAGATCTTTGTGACCATGCGCCATATATCCATGATGAGCTA 534  
RESULT 14  
AA97837  
ID AA97837 standard; cDNA: 832 BP.  
AC AA97837;  
XX  
XX 23-SEP-1999 (first entry)  
DE Human secreted protein encoding cDNA #25.  
XX  
XX Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;  
KW diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
XX  
OS Homo sapiens.  
XX  
XX W09925825-A2.  
PD 27-MAY-1999.  
XX  
XX 13-NOV-1998; 98WO-1B01862.  
PF 04-SEP-1998; 98US-0099273.  
XX 13-NOV-1997; 97US-0066677.  
PR 17-DEC-1997; 97US-0069957.  
PR 09-FEB-1998; 98US-0074121.  
PR 13-APR-1998; 98US-0081563.  
PR 10-AUG-1998; 98US-0096116.  
XX  
XX (GIST ) GENSET.  
XX  
XX Bouqueleret L, Duclet A, Dumas Milne Edwards J;  
PI WPI: 1999-347472/29.  
XX  
XX P-PSDB: AAY36153.  
DR  
XX  
XX Extended cDNAs encoding secreted proteins  
PT  
XX  
XX Example 28: Page 194-195; 307pp; English.  
PS  
XX  
XX AA97813-X97906 represent extended cDNA's which encode novel human  
CC secreted proteins (see AAY36129-136222) and which have cytosolic,  
CC thrombotic and osteopathic activity. The extended cDNAs can be used to  
CC express secreted proteins or parts of them or to obtain antibodies  
CC capable of binding to the secreted proteins. They may also be used in  
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC Uses also include design of expression vectors and secretion vectors.  
XX  
XX  
SQ Sequence 832 BP; 224 A; 194 C; 243 G; 168 T; 3 other:  
Alignment Scores:  
Pred. NO.: 1.5e-97 Length: 832  
Score: 911.00 Matches: 180  
Percent Similarity: 97.84% Conservative: 1  
Best Local Similarity: 97.30% Mismatches: 1  
Query Match: 94.90% Indels: 3  
Gaps: 0  
US-10-082-502-17 (1-182) x AA97837 (1-832)  
0Y 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyAlaLeuGlyThrAlaTrpAla 20  
|||||  
Db 148 ATGAAGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 207  
0Y 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGlu-LeuGluTr 40

Db 208 CGAGAGAGCCAGCATCTCCACTGTGAGCATGACGGCTGTGGATGAACATGAGATG 267  
0Y 40 pGluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAspR 60  
Db 268 GGAATATGCGCCAGGTGGACCCCAAGAGACCAATTCAGATGGATCTTCCGATCAATCC 327  
0Y 60 cAspLysSerGlnSerValAlaGluValProThrAlaArgSerGluAlaHisLeuThrG 80  
Db 328 AGATGGCAGCCAGTCAGTGGTGGAGTGCCTTATCCCGCTAGAGGCCCACTCACA 387  
0Y 80 uLeuLeuGluGluIleCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerTh 100  
Db 388 GCTCTGGAGAGATATGTGACCGGATGAGAGATGAGGAAACAGATTCATCTTCAC 447  
0Y 100 rHisArgLysAsnTrpValArgValAlaGlyArgAsnGlyGluSerSerGluLeuAspLe 120  
Db 448 CCATCGCAAGAACTACGTACGTGTAGTGGCCGGAATGAGATCCAGTGAACATGAGCT 507  
0Y 120 uGlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLys-PheAlaCysGluSerI 140  
Db 508 ACAAGGCATCCGAATCGACTGATATTAGGGCACCCCTCAAGBTTCGTGTGGAAACA 567  
0Y 140 ILeValGluGluTrpGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValL 160  
Db 568 TTGTGAGAGAAATACGAGATGAACTGAAATCTTTCCGAGAGGCTGCAATGTGA 627  
0Y 160 yAspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSer-115 179  
Db 628 AAGACAAACTTTGCACTGTAAGCAACAGATCTTTGACCAATGCCCTGCACATATCGCAT 687  
0Y 180 AspGluLeu 182  
Db 688 GATGAGCTA 696  
RESULT 15  
AAA43525  
ID AAA43525 standard; cDNA: 592 BP.  
XX  
XX AAA43525;  
AC  
XX  
XX 21-AUG-2000 (first entry)  
DE  
XX  
XX Mouse secreted expressed sequence tag SEQ ID NO:100.  
DE  
XX  
XX Human; mouse; chicken; rat; secreted expressed sequence tag; SPST;  
KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
KW immunomodulatory; haematopoietic; chemokinetin; analgesic; haemostatic;  
KW thrombolytic; antiinflammatory; cytosolic; antibacterial; antifungal;  
KW antiviral; antidiabetic; antisthmatic; vulnerrary; antiparkinsonian;  
KW anticancer; osteopathic; neuroprotective; noctropic; antiparasitic;  
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
KW central nervous system disorder; Alzheimer's disease; stroke;  
KW Parkinson's disease; Huntington's disease; coagulation disorder;  
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
KW tumour; infection; depression; psoriasis; ss.  
XX  
XX  
XX Mus musculus.  
OS  
XX  
XX W0200021991-A1.  
PN  
XX  
XX 20-APR-2000.  
PD  
XX  
XX 15-OCT-1999; 99WO-US24206.  
PF  
XX  
XX 15-OCT-1998; 98US-0104436.  
PR  
XX  
XX (GENY ) GENETICS INST INC.  
PA  
XX  
XX Jacobs K, McCoy JM, Lavallic ER, Collins-Racie LA, Evans C;  
PI

PI Merberg D, Treacy M, Bowman MR;  
 XX WPI: 2000-317938/27.  
 PT Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (seSTs), useful for treating various disorders  
 XX such as autoimmune, infectious, and central nervous system disorders -  
 PS Claim 1: Page 228-229; 803pp; English.

CC AAA43426 to AAA45925 represent specifically claimed secreted expressed  
 CC sequence tags (seSTs), isolated from human, mouse, chicken and rat  
 CC tissue sources. The seSTs can have a range of activities depending on  
 CC the tissues they were isolated from. The activities include:  
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
 CC cytostatic; analgesic; haemostatic; thrombolytic; antiinflammatory;  
 CC antiasthmatic; vulnerrary; antitumor; osteoporotic; neuroprotective;  
 CC neurotrophic; antiparkinsonian; antipsoriatic; cerebroprotective;  
 CC anticonvulsant; and antidepressant. The seSTs can be used for gene  
 CC therapy and in vaccines. The seSTs are useful as probes for the  
 CC identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the seSTs. Proteins encoded by the seSTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
 CC in the exemplification of the present invention.

SO Sequence 592 BP; 159 A; 139 C; 167 G; 123 T; 4 other;

Alignment Scores:  
 Pred. No.: 7.2e-89 Length: 592  
 Score: 835.50 Matches: 165  
 Percent Similarity: 93.48% Conservative: 3  
 Best Local Similarity: 91.85% Mismatches: 10  
 Query Match: 87.03% Indels: 4  
 DB: 21 Gaps: 0

US-10-082-502-17 (1-182) x AAA43525 (1-592)

QY 1 MetTysGlyTrpGlyTrpLeuAla-LeuLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAl 20  
 DB 44 ATGAAAGGNTGGGGTGGCTAGCCCTACTTTGGGGGCTCTGCGGAACCTGCGGGA 103  
 QY 20 AATGATGSeGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTr 40  
 DB 104 TCGAAGAGGCCAANATCTACACTGTGGANCTTCAGAGGCTCTGTGGATGAATTAAGATG 163  
 QY 40 pGluTleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPr 60  
 DB 164 GGAATTTGCCCGGTGGACCCCAAGAACATTCAGATGGGATCCTTCGAATCAATCC 223  
 QY 60 oAspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrG 80  
 DB 224 AGATGGCACCACACTGCTGTGGAGTACCTTATGCCCGTCAGAGGCCACCTCAGACA 283  
 QY 80 uLeuLeuGluGluTleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerTh 100  
 DB 284 GTTGTCTTGGAGAGTGTGTGACCGAATGAAGAGTACGGGA-CAGATTGA-CCTTTCTAC 341  
 QY 100 rHisArgLysAsnTyrValArgValValGlyArgAsnGlyGluSerSerGluLeuAspLe 120  
 DB 342 CCACCGCAAGAACTACGTACGCGTGTGCGGATGAGATGCAATCAGTAAGTAAGTACT 401  
 QY 120 uGlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerI 140  
 DB 402 ACAGGGCAATCCGAATTGACTCAGATATATCANGCGACCCCTCAAGTTTGGTGTGAGAGCAT 461

QY 140 eValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValY 160  
 DB 462 TGTGGAAGAAATACGAGGATGAGCTTATCGAATTTCTTCCAGAGAGGCTGCAACGTTAA 521  
 QY 160 sAspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSer-H1SA 180  
 DB 522 AGACAACTTTGCAGTTAGCGGACAGATCTATGTGACCATCCCTCGACAGATCTCAGC 581  
 QY 180 sPGLuLeu 182  
 DB 582 AAGAGCTC 589

Search completed: March 30, 2003, 22:42:53  
 Job time: 145.849 secs



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:33:58 ; Search time 39.6802 Seconds

(without alignments)  
1406.624 Million cell updates/sec

Title: US-10-082-502-17

Perfect score: 960  
Sequence: 1 MKGGMWALLGALGATGA.....LCNRPTDCHALHSHDEL 182

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2\_1/USPRO.spool/US10082502/runat\_24032003\_135100\_6616/app-query.fasta.1.1308  
-DB=Issued\_Patents\_NA -QWMT=fastap -SOFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd  
-LIST=45 -DOCCALIGN=200 -THR\_SCORE=Pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10082502.ECGN.1.1.131 @runat\_24032003\_135100\_6616 -NCPU=6 -ICPU=3  
-NO\_XLIPY -NO\_MMAP -LARGOQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA:  
1: /cgn2\_6/ptodata/1/lna/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/lna/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/lna/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/lna/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/lna/PCPUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/lna/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	131.5	13.7	744	US-09-163-285-3	Sequence 3, Appl1
2	131.5	13.7	744	US-09-163-285-3	Sequence 1, Appl1
3	131.5	13.7	744	US-09-163-285-3	Sequence 14, Appl1
4	85.5	8.9	61663	US-09-453-7028-62	Sequence 62, Appl1
5	85.5	8.9	61663	US-08-257-073-10	Sequence 10, Appl1
6	84	8.6	3752	US-08-961-083-159	Sequence 159, Appl1
7	83	8.6	3752	US-08-961-527-208	Sequence 208, Appl1
8	81.5	8.5	1216	US-09-071-035-3	Sequence 3, Appl1
9	81.5	8.5	1367	US-09-071-035-3	Sequence 1, Appl1
10	81.5	8.5	4394	US-08-750-152A-1	Sequence 1, Appl1
11	81.5	8.5	4403765	US-09-103-840A-2	Sequence 2, Appl1
12	81.5	8.5	4411529	US-09-103-840A-1	Sequence 1, Appl1

13	81	8.4	833	2	US-08-837-029-1	Sequence 1, Appl1
14	79.5	8.3	1173	3 <td>US-08-993-380-3<td>Sequence 3, Appl1</td></td>	US-08-993-380-3 <td>Sequence 3, Appl1</td>	Sequence 3, Appl1
15	78.5	8.1	5661	4 <td>US-08-938-105-2<td>Sequence 2, Appl1</td></td>	US-08-938-105-2 <td>Sequence 2, Appl1</td>	Sequence 2, Appl1
16	78	8.2	1530	4 <td>US-09-131-831B-2<td>Sequence 2, Appl1</td></td>	US-09-131-831B-2 <td>Sequence 2, Appl1</td>	Sequence 2, Appl1
17	78	8.1	1554	3 <td>US-08-809-999D-1<td>Sequence 1, Appl1</td></td>	US-08-809-999D-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
18	78	8.1	1554	3 <td>US-09-069-637-1<td>Sequence 1, Appl1</td></td>	US-09-069-637-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
19	78	8.1	1554	4 <td>US-09-322-360-1<td>Sequence 1, Appl1</td></td>	US-09-322-360-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
20	78	8.1	1554	4 <td>US-09-131-831B-1<td>Sequence 1, Appl1</td></td>	US-09-131-831B-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
21	78	8.1	1555	2 <td>US-08-316-218B-1<td>Sequence 1, Appl1</td></td>	US-08-316-218B-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
22	78	8.1	2148	3 <td>US-08-809-999D-2<td>Sequence 2, Appl1</td></td>	US-08-809-999D-2 <td>Sequence 2, Appl1</td>	Sequence 2, Appl1
23	78	8.1	2148	3 <td>US-09-069-637-2<td>Sequence 2, Appl1</td></td>	US-09-069-637-2 <td>Sequence 2, Appl1</td>	Sequence 2, Appl1
24	78	8.1	2148	4 <td>US-09-322-360-2<td>Sequence 2, Appl1</td></td>	US-09-322-360-2 <td>Sequence 2, Appl1</td>	Sequence 2, Appl1
25	77.5	8.1	1776	2 <td>US-08-522-229B-1<td>Sequence 2, Appl1</td></td>	US-08-522-229B-1 <td>Sequence 2, Appl1</td>	Sequence 2, Appl1
26	77.5	8.1	1776	2 <td>US-09-102-433-1<td>Sequence 1, Appl1</td></td>	US-09-102-433-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
27	77.5	8.1	2508	4 <td>US-08-850-993-1<td>Sequence 1, Appl1</td></td>	US-08-850-993-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
28	77	8.0	3218	4 <td>US-09-221-017B-255<td>Sequence 255, Appl1</td></td>	US-09-221-017B-255 <td>Sequence 255, Appl1</td>	Sequence 255, Appl1
29	76.5	8.0	2012	2 <td>US-08-484-200-3<td>Sequence 3, Appl1</td></td>	US-08-484-200-3 <td>Sequence 3, Appl1</td>	Sequence 3, Appl1
30	76.5	8.0	2478	4 <td>US-08-485-355B-47<td>Sequence 47, Appl1</td></td>	US-08-485-355B-47 <td>Sequence 47, Appl1</td>	Sequence 47, Appl1
31	76.5	8.0	2478	4 <td>US-08-485-355B-49<td>Sequence 49, Appl1</td></td>	US-08-485-355B-49 <td>Sequence 49, Appl1</td>	Sequence 49, Appl1
32	76.5	8.0	2478	4 <td>US-09-194-613-1<td>Sequence 51, Appl1</td></td>	US-09-194-613-1 <td>Sequence 51, Appl1</td>	Sequence 51, Appl1
33	76.5	8.0	2479	4 <td>US-08-485-355B-51<td>Sequence 51, Appl1</td></td>	US-08-485-355B-51 <td>Sequence 51, Appl1</td>	Sequence 51, Appl1
34	76.5	8.0	4865	3 <td>US-08-894-017-24<td>Sequence 38, Appl1</td></td>	US-08-894-017-24 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
35	76	7.9	4131	1 <td>US-07-872-644-38<td>Sequence 38, Appl1</td></td>	US-07-872-644-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
36	76	7.9	4131	1 <td>US-08-297-494-38<td>Sequence 38, Appl1</td></td>	US-08-297-494-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
37	76	7.9	4131	1 <td>US-08-297-510-38<td>Sequence 38, Appl1</td></td>	US-08-297-510-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
38	76	7.9	4131	1 <td>US-08-479-532-38<td>Sequence 38, Appl1</td></td>	US-08-479-532-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
39	76	7.9	4131	1 <td>US-08-455-526-38<td>Sequence 38, Appl1</td></td>	US-08-455-526-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
40	76	7.9	4131	1 <td>US-08-455-526-38<td>Sequence 38, Appl1</td></td>	US-08-455-526-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
41	76	7.9	4131	1 <td>US-09-139-491-38<td>Sequence 38, Appl1</td></td>	US-09-139-491-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
42	76	7.9	4131	5 <td>PCT-US92-03222-38<td>Sequence 38, Appl1</td></td>	PCT-US92-03222-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
43	75	7.8	5558	4 <td>US-08-961-527-103<td>Sequence 103, Appl1</td></td>	US-08-961-527-103 <td>Sequence 103, Appl1</td>	Sequence 103, Appl1
44	75	7.8	6519	1 <td>US-08-588-985-1<td>Sequence 1, Appl1</td></td>	US-08-588-985-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
45	75	7.8	6519	1 <td>US-08-971-988-1<td>Sequence 1, Appl1</td></td>	US-08-971-988-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-163-285-3  
Sequence 3, Application US/09163285  
Patent No. 6204013

GENERAL INFORMATION:  
APPLICANT: Rhododoust, Mehran  
TITLE OF INVENTION: NOVEL MSP-5 PROTEIN AND NUCLEIC ACID MOLECULES  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHYE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/163,285  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/090,398  
FILING DATE: June 24, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragoras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MNI-049  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 3:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 744 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..744
US-09-163-285-3

Alignment Scores:
Pred. No.: 2,828-07 Length: 744
Score: 131.50 Matches: 50
Percent Similarity: 44.74% Conservative: 35
Best Local Similarity: 26.32% Mismatches: 62
Query Match: 13.70% Indels: 43
DB: Gaps: 11

US-10-082-502-17 (1-182) x US-09-163-285-3 (1-744)
Oy 7 LeuAlaLeuLeuGlyAlaLeuLeuGlyThr-----AlaTrpAla-----Arg 21
    ||| :|||:||||| ||| :|||:|||||
Db 16 TTGGGAATATTCCTTTCTTTTGGCCGCGACAGAGCCTTGCGCTGGATGTTGAAG 75
Oy 22 ArgSerGlnAsp-----LeuHisCysGlyAlaCysArgAlaLeuVal 35
    ||| :|||:||||| ||| :|||:|||||
Db 76 GAGGAGACGATGACACAGAACGCTTGCCCAACCAATGCGAAGTGTGTACCTGCTGAGC 135
Oy 36 AspGluLeuGluTrpGluLeuLeuGlnValAsp---ProValysThrIleGlnMetGly 54
    ||| :|||:||||| ||| :|||:|||||
Db 136 ACAGAGCTACAGAGCGGAACTGATGCGACCGCTCATCTCGAGAGAGTGTGAGTGGGG 195
Oy 55 SerPheArgIleAsnProAspGlySerGlnSerValValGluValProTyrAlaArgSer 74
    ||| :|||:||||| ||| :|||:|||||
Db 196 CAGGTGCTG-----GATACAGGCAAGAGAGAGACAGAGCTGCTTACAGCGCTTCA 246
Oy 75 GluAlaHisLeuThrGluLeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGlu 94
    ||| :|||:||||| ||| :|||:|||||
Db 247 GAGACAGAGCTGGAAGAGCCCTTAGAGAAATTATGTGAGCGGATCTCGAGACTAT----- 300
Oy 95 GlnIleAspProSerThrHis-----ArgLysAsnTyrValArgValValGlyArgAsn 112
    ||| :|||:||||| ||| :|||:|||||
Db 301 -----AGTGTTCACGCTGACGCGCAAGGCTCAGTACAGATATGACC-----AAG 342
Oy 113 GlyIleuSerSerGlu-----LeuAspLeuGlnGlyIleArgIleAsp 126
    ||| :|||:||||| ||| :|||:|||||
Db 343 GGTGAGAGTACAGACCATGGCAACACTGAAGGCTTAGTGCACAAAGGGGTGAAGTGGAT 402
Oy 127 SerAspIleSer-----GlyThrLeuLysPhe 135
    ||| :|||:||||| ||| :|||:|||||
Db 403 CTGGGAGATCCCTCTGAGACCTTGGGATGAGCCACGCTGAGAGTCAACATACCTCAAGAG 462
Oy 136 AlaCysGluSerIleValGluLeuTrpGluAspGluLeuIleGlu---PhePheSerArg 154
    ||| :|||:||||| ||| :|||:|||||
Db 463 CAGTGTGAGACCATGTTGGAGAGATTGAAAGCATTTGGGAGACTGGTACTTCCACCAT 522
Oy 155 GluAlaAspAsnValIleAspLysLeuGly 164
    ||| :|||:||||| ||| :|||:|||||
Db 523 CAGGAGCAGACCCCTACAAATAATTTCTCTGT 552

RESULT 2
US-09-163-285-1
; Sequence 1, Application US/09163285
; Patent No. 6204013
; GENERAL INFORMATION:
; APPLICANT: Rhododoust, Mehran
; TITLE OF INVENTION: NOVEL MSP-5 PROTEIN AND NUCLEIC ACID MOLECULES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street

```

```

CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/163,285
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/090,398
FILING DATE: June 24, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MNI-049
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1512 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 125..868
US-09-163-285-1

Alignment Scores:
Pred. No.: 7,848-07 Length: 1512
Score: 131.50 Matches: 50
Percent Similarity: 44.74% Conservative: 35
Best Local Similarity: 26.32% Mismatches: 62
Query Match: 13.70% Indels: 43
DB: Gaps: 11

US-10-082-502-17 (1-182) x US-09-163-285-1 (1-1512)
Oy 7 LeuAlaLeuLeuGlyAlaLeuLeuGlyThr-----AlaTrpAla-----Arg 21
    ||| :|||:||||| ||| :|||:|||||
Db 140 TTGGGAATATTCCTTTCTTTTGGCCGCGACAGAGCCTTGCGCTGGATGTTGAAG 199
Oy 22 ArgSerGlnAsp-----LeuHisCysGlyAlaCysArgAlaLeuVal 35
    ||| :|||:||||| ||| :|||:|||||
Db 200 GAGGAGACGATGACACAGAACGCTTGCCCAACCAATGCGAAGTGTGTACCTGCTGAGC 259
Oy 36 AspGluLeuGluTrpGluLeuLeuGlnValAsp---ProValysThrIleGlnMetGly 54
    ||| :|||:||||| ||| :|||:|||||
Db 260 ACAGAGCTACAGAGCGGAACTGATGCGACCGCTCATCTCGAGAGAGTGTGAGTGGGG 319
Oy 55 SerPheArgIleAsnProAspGlySerGlnSerValValGluValProTyrAlaArgSer 74
    ||| :|||:||||| ||| :|||:|||||
Db 320 CAGGTGCTG-----GATACAGGCAAGAGAGAGACAGAGCTGCTTACAGCGCTTCA 370
Oy 75 GluAlaHisLeuThrGluLeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGlu 94
    ||| :|||:||||| ||| :|||:|||||
Db 371 GAGACAGAGCTGGAAGAGCCCTTAGAGAAATTATGTGAGCGGATCTCGAGACTAT----- 424
Oy 95 GlnIleAspProSerThrHis-----ArgLysAsnTyrValArgValValGlyArgAsn 112
    ||| :|||:||||| ||| :|||:|||||
Db 425 -----AGTGTTCACGCTGAGCGCAAGGCTCAGTACAGATATGACC-----AAG 466
Oy 113 GlyIleuSerSerGlu-----LeuAspLeuGlnGlyIleArgIleAsp 126
    ||| :|||:||||| ||| :|||:|||||
Db 467 GGTGAGTACAGACCATGGCAACACTGAAGGCTTAGTGCACAAAGGGGTGAAGTGGAT 526

```



[illegible]

Query Match:	8.91%	Indels:	30
DB:	4	Gaps:	6

```

OY      5  GLLTTPLEALALEULEUENCLYLALEULEUGLYTHRALATRALARGTSGEIN 24
Db      37859  GGCTGGCTGCGGACGATTATACGAGTTTGTGTGGACCACTTGGGATCAACAGGGGGCAG 37918
OY      25  APLLEUHLSCYSGLYALACYSARGHLALEUVALAEPJULEUGLUTRPSLULE----- 42
Db      37919  ACAGTGCATTTGATCCCGTAAAAAAGTTTAAACGAGGTGACGTGACGTGCTGTGTGT 37978
OY      43  -----AAGLGNVAL-----AAPPLQVLSLYS-THRIIEGLMETCL 54
Db      37979  TTCGGGACGCTGGAAMATTCCTGTATGGTATGAGCGCATATACGAGCAAMATTTATTCCTGAC 38038
OY      54  YSERHEARTGLEAENPPOASPGLYSERGLNSERVALVALGVALPROTYALARGSE 74
Db      38039  GGCGGTGACAGTTTACCCGGATGCCCTTAGCGGACCACTGGCTGGCAGTACCCCGTGA 38098
OY      74  RLGLAHLASLEUTHRLGULEULEUGLUGLULECYASPARGMETLSGLUTRGLYGL 94
Db      38099  AGGTGCATAT-----TACGGAAAT 38116
OY      94  UGLINLEASPPOSERTHRHLSATGLYSASNTYRVALRVALVALGLYATGASNGLYGL 114
Db      38117  GACGCTGATGATCCGGGGGAGTATGCGGTTTCG---GTGACGGGTGAAGGAAAGAACACTGCTGT 38173
OY      114  USERERGLUEAUPLEUGLNGLYLLEARGLEASPERASPILESERGLYTHLEULY 134
Db      38174  CTACGGACGTGTGGCTTATGAGGGACCGCAAAAGTACGTGACGCTCAATATGCTGTTCG 38233
OY      134  SPHEALCYSGUSERILEVALGLU---GLUTYRGLUASPGLULEULEGLUPHEPHESE 153
Db      38234  C-----GCGACTCTCTGTTGTGACGTTTAAAGCATACCCGAGAACACTGCTGACAGATTTCG 38284
OY      153  FARGGLUHLAASPASVALYLSAPPLYSLEUCYSERLYSARG 167
Db      38285  CGACATACAGAAATATGTGGCTGATGACCTTGCACCTATTTCGT 38327

RESULT 5
US-08-257-073-10
: Sequence 10, Application US/08257073
: Patent No. 5766597
: GENERAL INFORMATION:
: APPLICANT: Paoletti, Enzo
: APPLICANT: de Talaene, Charles
: APPLICANT: Tine, John A.
: TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
: NUMBER OF SEQUENCES: 143
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Curtiss, Morris & Safford, P.C.
: STREET: 530 Fifth Avenue, 25th Floor
: CITY: New York
: STATE: New York
: COUNTRY: UNITED STATES OF AMERICA
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/257,073
: FILING DATE: 09-JUN-1994
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/075,783
: FILING DATE: 11-JUN-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/852,305
: FILING DATE: 18-MAR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/672,183
: FILING DATE: 20-MAR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Frommer, William S.

```

```

      .506 REGISTRATION NUMBER: 25,506
      REFERENCE/DOCKET NUMBER: 454310-2570
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (212) 840-3333
      TELEFAX: (212) 840-0712
      TELEX: 425066 CURTMS
      INFORMATION FOR SEO ID NO: 10:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 5181 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
US-08-257-073-10

Alignment Scores:
Pred. No.:          2.76           Length:       5181
Score:             85.00           Matches:     36
Percent Similarity: 43.31%        Conservative: 32
Best Local Similarity: 22.93%     Mismatches:  59
Query Match:       8.85%          Indels:      30
DB:               1              Gaps:        7

US-10-082-502-17 (1-182) x US-08-257-073-10 (1-5181)

OY   36 ASPEGLUENGUTRPGLIUEALGINVALASPPLYSLYTHTLIEINMELYSER 55
    ||| ||||| ||| ::: ||| ::: :::
Db   1192 GATCACCATTGAATAGATATTTATTAGCGAAAAAATAAAAGGTGATGCACCT 1251
OY   56 PHEATGILEANPPOASPGLYSERSGINSEVALVAILGVALPROTYRLAARGSERLU 75
    ::||| :::||| ||| ::: ::: ||||| |||
Db   1252 AAATCAACAAGTCCTACGAAATCTGTTCAAATRCCAAAAGTCTTATCCAANGCTATT 13111
OY   76 ALAHIS-----LEUTHRGULEUGLUGLIUECYA-----ASPARWELYS 90
    ::||| :|||||: ::: ::: ::: ::: :::
Db   1312 GTATATCCTTTACACACTACACTGATTAATCATTAATCTTGCTGCAGATAATGATTAANAAT 13711
OY   91 GLUTYLGLUGLINILLESPPROSETHRHSAIGLYASAHTYVALARYVALICY 110
    ||||| :::||| ||| ||| :::|||
Db   1372 TCATATGGTATTTAATGATCTCTGATCTAATAAGAAAAATTATGAAAAAATTATTACA 14311
OY   111 ARGANGLYGU-----SERSGULUENAPLEUNGILNLY 122
    ||| ||| ::: ::: ::: ::: ::: ::: :::
Db   1432 GATTAATAGCAAGAAAAAATTTCAATTATACATTAAAAAACAAATGATTTACAGAA 1491
OY   123 ILEARGLIESPERASPILESERGLYHYLEYSPHEALACYGLUSERILEVALICU 142
    ||||| :::||| ||| ::: ::: |||
Db   1492 AAAAAATTATATC-----ACAAAAAGCAAAATRAAAAAATTACTTGA 1536
OY   143 GLUTYGLU-----ASPEGLULEUILLGLU---PHEPSESERARGLU 155
    ::||| ||| ::: ::: ::: ::: ::: :::
Db   1537 GATTAATAGCAAGAAAAAAGATTTATGAAGAATTCCTGAAAAATTTATGCAAAATGAA 1596
OY   156 ALASPSAN-----VALLYSAPPYLSLEYCSERYLARG 167
    ::||| ||| ||| |||||: ::: ||| |||
Db   1597 TTTAATATAATTTTGACAAAGATGTCGATGAATAAATATTCAGTCAGAA 1647

RESULT 6
US-08-961-083-159
Sequence 159, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
```

917 CAGGATGAATTAGA

OY 126 AspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGluTyrGlu 145  
DB 977 CAAATTAAGTTCGTGATTTA-----CAAAAAGAA 1006  
OY 146 AspGluLeuIleGluPhePheSerArgGluAlaAsp-----AsnVal 159  
DB 1007 ATTAGTACCTGTAATATATTACTGGAGGGGCTGATCCTGAGAGATGATCTGCTCTT 1066  
OY 160 LysAspLysLeuLysSerLysArgThrAspLeu 170  
DB 1067 CAAATTAATGCTGCTTAATAAAGCTGAGTA 1099

## RESULT 8

US-09-071-035-3  
Sequence 3, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071.035  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1216 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-071-035-3

Alignment Scores:  
Pred. No.: 0.931 Length: 1216  
Score: 81.50 Matches: 41  
Percent Similarity: 41.81% Conservative: 33  
Best Local Similarity: 23.16% Mismatches: 58  
Query Match: 8.49% Indels: 45  
DB: 4 Gaps: 10

US-10-082-502-17 (1-182) x US-09-071-035-3 (1-1216)

OY 3 GYTTPGTYTTPLeuAlaLeuLeuGluAlaLeuLeuGlyThrAlaTyrAlaArgArg 22  
DB 260 GCGTGTATGGAATGCT----- 277  
OY 23 SerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGlu-----LeuGluTyr 40  
DB 278 GCACAAAGATGAATGTATGTGACTTAACACATATATGATGATGACACAAATCGGCTGG 337

OY 41 GluIleAlaGluValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 338 AAGATGCA-----GAGCCAAATTCGTGAAGTATGTTAGAGCGCCCAAAATC----- 385  
OY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisIleThrGlu 80  
DB 386 GACGGCAACAA-----TAGCGCATTCATTAATTAATTCG-----ACAGAA 427  
OY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
DB 428 ATGTATCTGTAAATAGCTGATTTGTTGAAGAATATGCGTTGAAGTA---CCGAAACAA 484  
OY 101 -----HisArgLysAsnTyrValArgValAl 109  
DB 485 TTAGAGCAATTAAGAACCTCTTAACAAATTAACGAAAAATCCACAAAGAGTCGTT 544  
OY 110 GYATGAsnGlyLysSerGlu-----LeuAspLeuGlnGlyIle 123  
DB 545 GGTGCTGTTTGACTCGTTAATTAATCTTACGCAATGGAATGAAAAACAAAGCGTT 604  
OY 124 ArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGlu 143  
DB 605 GATTTTAATTAAGACTTACATTTTACACAAAGATTCG---CAAGAGTCGTGACCTAT 661  
OY 144 TyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
DB 662 TACCGTGATGTATCGAACAGGTTACTTCCGCACAGAGTGTGATGATAAA 712

## RESULT 9

US-09-071-035-1  
Sequence 1, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071.035  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1347 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-071-035-1

Alignment Scores:  
Pred. No.: 1.08 Length: 1347  
Score: 81.50 Matches: 41





QY 149 -----11egluPheSerArgGluAlaAspAsnValLysAsp 161  
Db 4389643 GCCCATGACATCGCGGTGGTGTGCGACAGTTCACGGGGAGTGTG-----GAC 4389596  
QY 162 LysLeuCySerLysArgThrAspLeuCyAspHisAla 174  
Db 4389595 GCGTGTGCTGCGAGATCATCCCGGTGGTGTGCGACGCC 4389557  
RESULT 13  
US-08-837-029-1  
Sequence 1, Application US/08837029  
Patent No. 5945303  
GENERAL INFORMATION:  
APPLICANT: Wei et al.  
TITLE OF INVENTION: Human Hematopoietic - Specific Protein  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: MD  
COUNTRY: US  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/837,029  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US96/04930  
FILING DATE: 11-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Brooks, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PF268  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 833 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 42..608  
FEATURE:  
NAME/KEY: s19.peptide  
LOCATION: 42..107  
FEATURE:  
NAME/KEY: mat.peptide  
LOCATION: 108..608  
US-08-837-029-1  
Alignment Scores:  
Pred. No.: 0.623 Length: 833  
Score: 81.00 Matches: 46  
Percent Similarity: 37.81% Conservative: 30  
Best Local Similarity: 22.89% Mismatches: 81  
Query Match: 8.44% Indels: 44  
Gaps: 10  
US-10-082-502-17 (1-182) x US-08-837-029-1 (1-833)  
QY 7 LeuAlaLeuLeuGluAlaLeu-----LeuGly----- 16  
Db 63 CTGCTGCTGCTGCTGCGAGCTGGCCATCCAGGGGGGCTGGGGGACAGGCCCATCTC 122

QY 17 ThrAlaThrAlaArgSerGlnAsp----- 25  
Db 123 ACAGCCACACGCCCACTGATGATGAGGATGCTACAGCCCATATGCCCGCTCAC 182  
QY 26 LeuHisCySerGlyAlaCySerArgAlaLeuValAspGluLeuGluTrpGluIleAlaGlnVal 45  
Db 183 CTGCCCTGTGATGCTGCGAGAGCTGTGCTTACAGATGTGGCAAAATCTGGCAAAAGCA 242  
QY 46 AspProLysLysThrIleGlnMetGlySerPheArgIleAsnProAspGlySerGlnSer 65  
Db 243 GAGACCAAACTCATACCTCA-----AACTTGGGGGGCGCGGAA 284  
QY 66 ValValGluValProTyrAlaArgSerGluAlaHisLeuThrGluLeuGluGluIle 85  
Db 285 CTGACCGAGTTGGTCTAC-----ACGATGTCTTGACCGGAC 323  
QY 86 CysAspArg---MetLysGluTyrGly---GluGlnIleAspProSerThrHisArgLys 103  
Db 324 TGCTCCCGAACTGCGAGACTACGAGGTTGCAAGATGAC----- 365  
QY 104 AsnTyrValArgValValGlyArgAsn---GlyGluSerSerGluLeuAspLeuGlnGly 122  
Db 366 CAACTGAACCTCTCACAGGCGCCAGACTTACGAGGGGCCAGACCAATCAGCTGCGTG 425  
QY 123 IleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCySerGluSerIleValGlu 142  
Db 426 ATGCTCACAGGGGGCCCTGCTGCTACCGAGCTCTCCAGACATGTTGCACACTTGGGG 485  
QY 143 GluTyr---GluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLysAsp 161  
Db 486 GAGTTTGGAGAGACAGATCTATGAAAGCCACCAACAGAGGGGCTTGGAGGCA 545  
QY 162 LysLeuCySerLysArgThrAspLeuCyAspHisAlaLeuHisIleSerHisAspGlu 181  
Db 546 TTGCTATGTGGGGAGCCCAAGGGGCTGCTCAGAGAGGTGTCCAGCACACAGAGAGAG 605  
QY 182 Leu 182  
Db 606 CTC 608  
RESULT 14  
US-08-993-380-3  
Sequence 3, Application US/08993380B  
Patent No. 6077676  
GENERAL INFORMATION:  
APPLICANT: Shi, Qianwei  
TITLE OF INVENTION: SINGLE CHAIN POLYPEPTIDES COMPRISING TROPONIN I AND  
FILE REFERENCE: 1112-1-053  
CURRENT APPLICATION NUMBER: US/08/993,380B  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 1173  
TYPE: DNA  
ORGANISM: Homo sapiens (modified)  
US-08-993-380-3  
Alignment Scores:  
Pred. No.: 1.57 Length: 1173  
Score: 79.50 Matches: 47  
Percent Similarity: 35.41% Conservative: 27  
Best Local Similarity: 22.49% Mismatches: 58  
Query Match: 8.28% Indels: 77  
Gaps: 8  
US-10-082-502-17 (1-182) x US-08-993-380-3 (1-1173)  
QY 10 LeuLeuGluAlaLeuLeuGluThrAlaThrAlaArgArgSerGlnAspLeuHisCySerGly 29  
Db 10 LeuLeuGluAlaLeuLeuGluThrAlaThrAlaArgArgSerGlnAspLeuHisCySerGly 29

```

Db 460 ATGATGACAGGGCTGCTGGG---GCCCGGGCTAAGAGACTCCCTGAGCTG----- 507
Qy 30 AlaCysArgAlaLeuValAspIleuGluIleuValAlaGlnValAspProIleu 49
Db 508 -----CGGGCCACCCTCAAGCAGGTGAAGAGAGACACCGAG-----AAGGAA 552
Qy 50 ThrIleGlnMetGlySerPheArgIleAsnProAsp----- 61
Db 553 AACCGGAGGTGGAGACTGGCGCAAGAACATGATGACATGAGTGAATGAGAGGCCGC 612
Qy 61 ----- 61
Db 613 AAGAAAGTTTGAGAGCAGCTAGTGGTGGTGGTGGTGGTGGGCGGCTTCTGGTGC 672
Qy 62 GlySerGlnSerValValGlnValProIleuValArgSerGlnAlaHisLeuThrGluLeu 81
Db 673 GGTGGTTTCATCATGATGATGATCATCTACAGAGCTGCGGTAGACAGCAGACAGAG 732
Qy 82 -----LeuGluGluIleCys 86
Db 733 CAGAAATGACTTCAAGGACGCTTCGACATCTTCGTCGCGGCGGTAGAGTGGCTGC 792
Qy 87 AsparGlnMetGluIleuValGluIleuValAspProSerThrIleArgLysAsnIleVal 106
Db 793 ATCAGCACCAGAGAGCTGGGCAAG-----GTGATG 822
Qy 107 ArgValValGlyArgAsnGlyIleuSerGlnLeuAspLeuGlnGlyIleArgIleAsp 126
Db 833 AGGATGCTGGGGCGAGACCCGCTGAGGAGCTGCAGAGATGATGATGATGATGATG 882
Qy 127 SerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleVal----- 141
Db 883 GAGGACGCGCAGCGGACGCTGACCTTTCAT---GAGTTCCTGGTATGATGATGCTGGCTGC 939
Qy 142 -----GluGluIleuValGluIleuValGluIleuValGluIleuValGluIleuVal 152
Db 940 ATGAGAGACGACAGCAAGGAAATCTGAGAGAGAGAGCTGTGACCTCTTCGCAATGTTT 999
Qy 153 SerArgGluAlaAspAsnValLysAsp 161
Db 1000 GACAAATGCTGATGGCTACATCGAC 1026

RESULT 15
US-08-938-105-2
Sequence 2, Application US/08938105
Patent No. 6353151
GENERAL INFORMATION:
APPLICANT: Leinwand, Leslie A.
APPLICANT: Vikstrom, Karen L.
TITLE OF INVENTION: TRANSGENIC MODEL FOR HEART FAILURE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Sheridan Ross P.C.
STREET: 1700 Lincoln St., Suite 3500
CITY: Denver
STATE: CO
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938.105
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crook, Wamell M.
REGISTRATION NUMBER: 31,071
REFERENCE/DOCKET NUMBER: 3595-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700

```

```

; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5661 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..5661
US-08-938-105-2

Alignment Scores:
Pred. No.: 20.2 Length: 5661
Score: 78.50 Matches: 37
Percent Similarity: 43.02% Conservative: 40
Best Local Similarity: 20.67% Mismatches: 71
Query Match: 8.18% Indels: 31
DB: Gaps: 7

US-10-082-502-17 (1-182) x US-08-938-105-2 (1-5661)
Qy 17 ThrAlaIlePheAlaArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAsp 36
Db 2656 ACGGCCAAGAGAGCGGAGCTGGAGAGC---GAGTGTCTAGAGCTCAAGAAAGATTCGAT 2712
Qy 37 GluLeuGluIlePheGlnValAspProLysValThrIleGlnMetGlySerPhe 56
Db 2713 GACCTGAGAGCTGACCTGCGCAAGGTGGAGAAAGCAAGCAAGCAAGCAAGT 2772
Qy 57 ArgIleAsnProAspGlySerGlnSerValValGluVal-----ProIleAspSer 74
Db 2773 AAAAAGCTGACAGAGAGATGGCCGCGCTGAGCAGAGATCATTCGCAAGCTGCACCAAG 2832
Qy 75 GluAlaHisLeuThrGluLeuGluGluIleCysAspArgMetLysGluIleValGlu 94
Db 2833 AAGAAAGCTTTCAGAGAGGCGCCAGCAGAGCCCTAGATGACCTTCAGGCTGAGAGAGC 2892
Qy 95 GlnIleAspProSerThrIleArgLysAsnIleValArgValValGlyArgAsnGlyIle 114
Db 2893 AAGGTCAACACACTGACC-----AAGTCTAAGTCAAGCTG-----GAG 2931
Qy 115 SerSerGluLeuAspLeuGlnGly-----IleArgIleAspSer 127
Db 2932 CAGCAGGTGGATGATCTGGAGGATCCCTGAGCAGCAGAGAGAGAGTGGCATGACCTG 2991
Qy 128 Asp-----IleSerGlyThrLeuLysPheAlaCysGluSerIleValGlu 142
Db 2992 GAGCGAGCAAGCGAGAGCTGAGAGGTGACCTGAAAGCTGACCCAGAGAGATCAGAGC 3051
Qy 143 GluIleGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLysAspLys 162
Db 3052 CTGGAAGACAGCAAGCTT-----CAGCTGAGGAGAAAG 3084
Qy 163 LeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAspLeu 181
Db 3085 CTCAGAGAGAAAGAGTTTGACATCAGTCAGCAGAAACAGTAAATAGAGAGCAGCAG 3141

Search completed: March 31, 2003, 04:04:16
Job time : 1662.66 secs

```



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 07:02:46 ; Search time 55.0233 Seconds

(without alignments)  
2814.757 Million cell updates/sec

Title: US-10-082-502-17

Perfect score: 960  
Sequence: 1 MKGGMALALGALGTAMA.....LCSKRDLDHALHSHDEL 182

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame\_p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool/US10082502/runat\_24032003.135103.6870/app.query.fasta\_1.1308  
-DB=published.Applications\_NA -QFMT=fastcap -SUFFIX=trpb -MIMATCH=0.1  
-LOOPT=0 -LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsnum62  
-TRANS=human40.csl -LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100  
-TRN\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NOR=ext -HEATSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10082502.6CCN\_1.1.145.6runat\_24032003.135103.6870  
-NCPU=6 -ICPU=3 -NO.XLPXY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEROUT=120 -WARN\_TIMEROUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications\_NA:\*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/2/pubpna/PT_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	960	100.0	814	10	US-09-864-864-309
2	960	100.0	1210	9	US-10-028-072-173
3	960	100.0	1210	9	US-10-121-049-173
4	960	100.0	1210	9	US-10-123-904-173

5	960	100.0	1210	9	US-10-140-470-173	Sequence 173, App
6	960	100.0	1210	9	US-10-175-746-173	Sequence 173, App
7	960	100.0	1210	9	US-10-176-918-173	Sequence 173, App
8	960	100.0	1210	9	US-10-176-921-173	Sequence 173, App
9	960	100.0	1210	9	US-10-137-865-173	Sequence 173, App
10	960	100.0	1210	9	US-10-140-474-173	Sequence 173, App
11	960	100.0	1210	9	US-10-142-831-173	Sequence 173, App
12	960	100.0	1210	9	US-10-143-011-173	Sequence 173, App
13	960	100.0	1210	9	US-10-140-002-173	Sequence 173, App
14	960	100.0	1210	9	US-10-142-819-173	Sequence 173, App
15	960	100.0	1210	9	US-10-123-262-173	Sequence 173, App
16	960	100.0	1210	9	US-10-142-823-173	Sequence 173, App
17	960	100.0	1210	9	US-10-121-050-173	Sequence 173, App
18	960	100.0	1210	9	US-10-141-755-173	Sequence 173, App
19	914	95.2	657	9	US-09-822-846-398	Sequence 398, App
20	266	27.7	564	10	US-09-864-864-228	Sequence 228, App
21	131.5	13.7	1457	10	US-09-846-573B-10	Sequence 10, App1
22	131.5	13.7	1470	9	US-10-036-041-58	Sequence 58, App1
23	131.5	13.7	1470	9	US-10-035-855-58	Sequence 58, App1
24	131.5	13.7	1470	9	US-10-174-590-479	Sequence 479, App
25	131.5	13.7	1470	9	US-10-176-758-479	Sequence 479, App
26	131.5	13.7	1470	9	US-10-175-737-479	Sequence 479, App
27	131.5	13.7	1470	9	US-10-175-737-479	Sequence 479, App
28	131.5	13.7	1470	9	US-10-175-737-479	Sequence 479, App
29	131.5	13.7	1470	9	US-10-175-737-479	Sequence 479, App
30	131.5	13.7	1470	9	US-10-176-482-479	Sequence 479, App
31	131.5	13.7	1470	9	US-10-176-757-479	Sequence 479, App
32	131.5	13.7	1470	9	US-10-176-757-479	Sequence 479, App
33	131.5	13.7	1470	9	US-10-176-913-479	Sequence 479, App
34	131.5	13.7	1470	9	US-10-180-552-479	Sequence 479, App
35	131.5	13.7	1470	9	US-10-180-552-479	Sequence 479, App
36	131.5	13.7	1470	9	US-09-931-836-58	Sequence 58, App1
37	131.5	13.7	1470	9	US-10-173-700-479	Sequence 479, App
38	131.5	13.7	1470	9	US-10-174-572-479	Sequence 479, App
39	131.5	13.7	1470	9	US-10-174-582-479	Sequence 479, App
40	131.5	13.7	1470	9	US-10-174-582-479	Sequence 479, App
41	131.5	13.7	1470	9	US-10-175-739-479	Sequence 479, App
42	131.5	13.7	1470	9	US-10-175-739-479	Sequence 479, App
43	131.5	13.7	1470	9	US-10-175-740-479	Sequence 479, App
44	131.5	13.7	1470	9	US-10-175-743-479	Sequence 479, App
45	131.5	13.7	1470	9	US-10-176-488-479	Sequence 479, App

#### ALIGNMENTS

RESULT 1  
US-09-864-864-309  
Sequence 309, Application US/09864864  
Patent No. US20020102679A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Secrist, Heather  
APPLICANT: Lodes, Michael J.  
APPLICANT: Algate, Paul A.  
APPLICANT: Fling, Steve P.  
APPLICANT: Mannion, Jane  
APPLICANT: Benson, Darin R.  
TITLE OF INVENTION: CARTER, DARRICK  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.523  
CURRENT APPLICATION NUMBER: US/09/864,864  
CURRENT FILING DATE: 2001-05-23  
NUMBER OF SEQ ID NOS: 341  
SOFTWARE: Corlita Invention Disclosure Database  
SEQ ID NO 309  
LENGTH: 814  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-864-864-309

**Alignment Scores:**

Pred. No.:	1,486-113	Length:	814
Score:	960.00	Matches:	18
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-10-082-502-17 (1-182) x US-09-864-864-309 (1-814)

QY	1	MetuyslylrrtprgylrrprrpleuAlaLeuLeuLeuLylAlaLeuLeuLylrrAlaTrpAla	20
Db	145	ATGAAAGCGTGGGGTGTGGCTGGCCCTCTTCTGGGGGCGCTGCTGGGAACCGCTGGGGCT	204
QY	21	ArgrrgserrGlnAspLeuHlaScyGlyAlaAcyrGAlaLeuValAspGluLeuGlnTrp	40
Db	205	CGGAGGAGCCAGGAGTTCrCACrTGrGGAGrATrGCAGGGCTCTGGTGGATrGAACrTGAATrGC	264
QY	41	GluIleuLaGluValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro	60
Db	265	GAATTTCGCCAGTGGAGACCCGMAAGAACATTTCAGATGGGATTTTCCGGATCAATCCA	324
QY	61	AspGlySerGlnSerValValGluValProTryAlaArgSerGluAlaHisLeuThrGlu	80
Db	325	GATGGAGCCAGCTCAGTGGTGGAGGTCCrTTCrCCCGCTCAGAGGCCrCACrTTCACAGAG	384
QY	81	LeuLeuGluGluIleCyAsnArgMetLysGluThrGlyGlnIleAsnProSerThr	100
Db	385	CTGTGGAGCGAGATrATrTGACCGGATrAAGAGrATrGGGAACrGATrTGAATrCTTCCACC	444
QY	101	HisArgLysAsnTryValArgValValGlyArgAsnGlyGlnSerSerGluLeuAspLeu	120
Db	445	CATCGCAAGAACrTACrTACrGTACrGTACrGTGGCGGAArTGGAGATrCAATrCAATrGAGCTTA	504
QY	121	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle	140
Db	505	CAGGCArTCCGAACrCGACrTAGATrATTAGCGGCACCTCAAGrTTCGGTGTGAACrGATT	564
QY	141	ValGluGluLysTryGlnAspGluLeuIleGluPhePheSerArgLysAlaAspAsnValLys	160
Db	565	GTGGAGGAATrACGAGATrTGAACrTCAATrTCTTTTCCGAGAGGCTGCAATrGTTTAAA	624
QY	161	AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisasp	180
Db	625	GACAAACTTTCAGTAAAGCAACrAGATrCTTGTGACATrGCCTGACrATrTGCATrGAT	684
QY	181	GluLeu 182	
Db	685	GAGCTA 690	
RESULT 2			
US-10-028-072-173			
: Sequence 173, Application US/10028072			
: Publication NO. US20030004311A1			
: GENERAL INFORMATION:			
:	APPLICANT:	Baker, Kevin P.	
:	APPLICANT:	Beresini, Maureen	
:	APPLICANT:	DeForge, Laura	
:	APPLICANT:	Desnoyers, Luc	
:	APPLICANT:	Elivartoff, Ellen	
:	APPLICANT:	Gao, Wei-Qiang	
:	APPLICANT:	Gerritsen, Mary E.	
:	APPLICANT:	Goddard, Audrey	
:	APPLICANT:	Godowski, Paul J.	
:	APPLICANT:	Gurney, Austin L.	
:	APPLICANT:	Sherwood, Steven	
:	APPLICANT:	Smith, Victoria	
:	APPLICANT:	Stewart, Timothy A.	
:	APPLICANT:	Tumas, Daniel	
:	APPLICANT:	Watanabe, Colin K	
:	APPLICANT:	Wood, William	
:	APPLICANT:	Zhang	

1	TITLE OF INVENTION:	
2	FILE REFERENCE:	
3	CURRENT APPLICATION NUMBER:	US/10/028, 072
4	CURRENT FILING DATE:	2001-12-19
5	PRIOR APPLICATION NUMBER:	60/049911
6	PRIOR FILING DATE:	1997-06-18
7	PRIOR APPLICATION NUMBER:	60/056974
8	PRIOR FILING DATE:	1997-08-26
9	PRIOR APPLICATION NUMBER:	60/059113
10	PRIOR FILING DATE:	1997-03-17
11	PRIOR APPLICATION NUMBER:	60/059115
12	PRIOR FILING DATE:	1997-09-17
13	PRIOR APPLICATION NUMBER:	60/059117
14	PRIOR FILING DATE:	1997-09-17
15	PRIOR APPLICATION NUMBER:	60/059122
16	PRIOR FILING DATE:	1997-03-17
17	PRIOR APPLICATION NUMBER:	60/059184
18	PRIOR FILING DATE:	1997-09-17
19	PRIOR APPLICATION NUMBER:	60/059363
20	PRIOR FILING DATE:	1997-09-18
21	PRIOR APPLICATION NUMBER:	60/059352
22	PRIOR FILING DATE:	1997-09-19
23	PRIOR APPLICATION NUMBER:	60/059588
24	PRIOR FILING DATE:	1997-09-19
25	PRIOR APPLICATION NUMBER:	60/059836
26	PRIOR FILING DATE:	1997-03-24
27	PRIOR APPLICATION NUMBER:	60/062350
28	PRIOR FILING DATE:	1997-10-17
29	PRIOR APPLICATION NUMBER:	60/062285
30	PRIOR FILING DATE:	1997-10-17
31	PRIOR APPLICATION NUMBER:	60/062287
32	PRIOR FILING DATE:	1997-10-17
33	PRIOR APPLICATION NUMBER:	60/062814
34	PRIOR FILING DATE:	1997-10-24
35	PRIOR APPLICATION NUMBER:	60/062816
36	PRIOR FILING DATE:	1997-10-24
37	PRIOR APPLICATION NUMBER:	60/063045
38	PRIOR FILING DATE:	1997-10-24
39	PRIOR APPLICATION NUMBER:	60/063082
40	PRIOR FILING DATE:	1997-10-31
41	PRIOR APPLICATION NUMBER:	60/063127
42	PRIOR FILING DATE:	1997-10-24
43	PRIOR APPLICATION NUMBER:	60/063327
44	PRIOR FILING DATE:	1997-10-27
45	PRIOR APPLICATION NUMBER:	60/063329
46	PRIOR FILING DATE:	1997-10-27
47	PRIOR APPLICATION NUMBER:	60/063550
48	PRIOR FILING DATE:	1997-10-28
49	PRIOR APPLICATION NUMBER:	60/063561
50	PRIOR FILING DATE:	1997-10-28
51	PRIOR APPLICATION NUMBER:	60/063704
52	PRIOR FILING DATE:	1997-10-29
53	PRIOR APPLICATION NUMBER:	60/063733
54	PRIOR FILING DATE:	1997-10-29
55	PRIOR APPLICATION NUMBER:	60/063735
56	PRIOR FILING DATE:	1997-10-29
57	PRIOR APPLICATION NUMBER:	60/063738
58	PRIOR FILING DATE:	1997-10-29
59	PRIOR APPLICATION NUMBER:	60/063755
60	PRIOR FILING DATE:	1997-10-17
61	PRIOR APPLICATION NUMBER:	60/064248
62	PRIOR FILING DATE:	1997-11-03
63	PRIOR APPLICATION NUMBER:	60/064809
64	PRIOR FILING DATE:	1997-11-07
65	PRIOR APPLICATION NUMBER:	60/065186
66	PRIOR FILING DATE:	1997-11-12
67	PRIOR APPLICATION NUMBER:	60/065846
68	PRIOR FILING DATE:	1997-11-17
69	PRIOR APPLICATION NUMBER:	60/066364
70	PRIOR FILING DATE:	1997-11-21
71	PRIOR APPLICATION NUMBER:	60/066453
72	PRIOR FILING DATE:	1997-11-24
73	PRIOR APPLICATION NUMBER:	60/066511

;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/066770  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/069212  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069228  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069334  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069694  
;; PRIOR FILING DATE: 1997-12-16  
;; PRIOR APPLICATION NUMBER: 60/072320  
;; PRIOR FILING DATE: 1998-01-23  
;; PRIOR APPLICATION NUMBER: 60/073612  
;; PRIOR FILING DATE: 1998-02-04  
;; PRIOR APPLICATION NUMBER: 60/074086  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/074092  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/077791  
;; PRIOR FILING DATE: 1998-03-12  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/079663  
;; PRIOR FILING DATE: 1998-02-27  
;; PRIOR APPLICATION NUMBER: 60/079728  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/080165  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081695  
;; PRIOR FILING DATE: 1998-04-14  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081818  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082999  
;; PRIOR FILING DATE: 1998-04-24  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085149  
;; PRIOR FILING DATE: 1998-05-12  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/086414  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/086430  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/087106  
;; PRIOR FILING DATE: 1998-05-28

;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088730  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088741  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090538  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07

Alignment Scores:  
Pred. No.: 2,54e-113 Length: 1210  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00 Conservative: 0  
Best Local Similarity: 100.00 Mismatches: 0  
Query Match: 100.00 Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-17 (1-182) x US-10-028-072-173 (1-1210)

QY 1 MetLysGLYTrpGLYTrpLeuAlaLeuLeuGLYAlaLeuLeuGLYThrAlaTrpAla 20  
DB 543 ATGAAGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 602

QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspLeuLeuGluTrp 40  
DB 603 CGAGAGAGCCAGGATGCTCCAGTGTGAGAGCTGAGGCTGTGTGATGATGACTGAGATGG 662

QY 41 GluLeuAlaGlnValAspProLysIsthrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 663 GAAATTTGCCAGGTGAGCCCAAGCAACATTGAGATGAGATGATTCGATTCATCA 722

QY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80  
DB 723 GATGCGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782

QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
DB 783 CTGCTGAGAGAGATATGTGACCGGATGAAGAGATGGGAAACATATGATCTTCCACC 842

QY 101 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyGluSerSerGluLeuAspLeu 120  
DB 843 CATGCCAAGAACTACGTACGTGTACTGTGCGCGAATGAGATCCAGTAACGTGACCTA 902

QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 903 CAAGCATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 962



QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyr 40  
Db 603 CGGAGGAGCCAGGATGCTCCACTGTGAGAGCGGCTTGCTGTGATGAACTAGAAATGG 662  
QY 41 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAspPro 60  
Db 663 GAAATTGCCAGGTGGAGCCCAAGAAACCATTCAGATGGGATCTTCCGGATCATATCCA 722  
QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
Db 723 GATGGCAGCAGTACGTGGTGGAGTGGCTTATGCCCGCTCAGAGGCCACCTCACAGAG 782  
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
Db 783 CTGCTGGAGAGATATGTGACCGGATGAAGAGATATGGGGAACAGATTGATCTCTCCACC 842  
QY 101 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerSerGluLeuAspLeu 120  
Db 843 CATCGCAAGACTACGTACGTATAGTGGCGCGAATGGAAATCCACTGAACTGGACCTA 902  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 903 CAAGGCATCCGAAATCAGTCACTAGATATTAGCGCGACCTCAAGTTGGCTGGAGAGCAT 962  
QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
Db 963 GTGGAGGAATACGAGATGTAACATCTTAATCTTTCCGAGAGGCTGCAATGTGTAAA 1022  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180  
Db 1023 GACAAACTTTCAGTAAAGCAACAGATCTTTGTGACCATCCCTGCACATATGCGATGAT 1082  
QY 181 GluLeu 182  
Db 1083 GAGCTA 1088

RESULT 5  
US-10-140-470-173  
; Sequence 173, Application US/10140470  
; Publication No. US2003002231A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C160  
; CURRENT APPLICATION NUMBER: US/10/140,470  
; CURRENT FILING DATE: 2002-05-06  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-140-470-173

Alignment Scores: 2.54e-113 Length: 1210  
Pred. No.: 2.54e-113

Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0  
US-10-082-502-17 (1-182) x US-10-140-470-173 (1-1210)

QY 1 MetLysGlyTyrPglTyrPLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTrrPa 20  
Db 543 ATGAAAGCGTGGGTTGGCTGGCCCTCTTCTGGGGGCCCTGCTGGAAACCGCTGGGCT 602  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyr 40  
Db 603 CGGAGGAGCCAGGATGCTCCACTGTGAGAGCGGCTTGCTGTGATGAACTAGAAATGG 662  
QY 41 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAspPro 60  
Db 663 GAAATTGCCAGGTGGAGCCCAAGAAACCATTCAGATGGGATCTTCCGGATCATATCCA 722  
QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
Db 723 GATGGCAGCAGTACGTGGTGGAGTGGCTTATGCCCGCTCAGAGGCCACCTCACAGAG 782  
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
Db 783 CTGCTGGAGAGATATGTGACCGGATGAAGAGATATGGGGAACAGATTGATCTCTCCACC 842  
QY 101 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerSerGluLeuAspLeu 120  
Db 843 CATCGCAAGACTACGTACGTATAGTGGCGCGAATGGAAATCCACTGAACTGGACCTA 902  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 903 CAAGGCATCCGAAATCAGTCACTAGATATTAGCGCGACCTCAAGTTGGCTGGAGAGCAT 962  
QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
Db 963 GTGGAGGAATACGAGATGTAACATCTTAATCTTTCCGAGAGGCTGCAATGTGTAAA 1022  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180  
Db 1023 GACAAACTTTCAGTAAAGCAACAGATCTTTGTGACCATCCCTGCACATATGCGATGAT 1082  
QY 181 GluLeu 182  
Db 1083 GAGCTA 1088

RESULT 6  
US-10-175-746-173  
; Sequence 173, Application US/10175746  
; Publication No. US20030027270A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C353

```

; CURRENT APPLICATION NUMBER: US/10/175,746
; PRIORITY FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-746-173

Alignment Scores:
Pred. No.: 2,54e-113 Length: 1210
Score: 960.00 Matches: 182
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-10-082-502-17 (1-182) x US-10-175-746-173 (1-1210)

QY 1 MetlysglyttrpglyttrpaulaleuleuleuglyalaleuleuglythrAlatrrpAla 20
Db 543 ATGAAGGCTGGGGTGGCTGGCCCTGCTGGGGGCGCTGGGAACGGCTGGGCT 602

QY 21 ArgatgserrginaspleuHiscysglyalacysatrgalaleuValaspluencuutrrp 40
Db 603 CGAGAGACCCAGAGATCTGACCTGAGACATGAGGGCTGTGGATGAATGAACTGAATGG 662

QY 41 GluilealaglnValasprolyasrthrilegmetglsyrPhearglleaspro 60
Db 663 GAAATGGCCAGGTGGAGCCCAAGAGACATTCAGATGGGATCTTCCGATCAATCCA 722

QY 61 AspglyserginserValalgluValprotyrAlaargserglnAlaHlsleuthrglu 80
Db 723 GATGACACCGAGTCAGTGTGGAGGTGCTTATGCCCTCAGAGGCCACCTCAGAG 782

QY 81 LeuleugluGluilecysAspargmetlysglyutryrglyuglnIleasproserthr 100
Db 783 CTGCTGGAGAGATATGTGACCGGATGAAGAGATGGGAAACAAATGATCTTCCACC 842

QY 101 HlsarglyAsanTyrrValargValalgluValargannglyluserSerGluDeuaspLeu 120
Db 843 CATGCCAAGAACTACGTACGTGTAGTGGCGCGGAATGAGAAATCCAGTAATGAGACCTA 902

QY 121 GlnnglyllearglleaspserraspIleserGlythrleuLysPheAlaCysgluserlle 140
Db 903 CAAGGCATCCCAATCGATCAGATATATGACGCGACCCCTCAAGTTGCGGTGAGAGCAT 962

QY 141 ValglugluTyrrGluAspGluLeuilegluphPheSerargGluAlaAspAsnVallys 160
Db 963 GTGGAGGAATACGAGAGATGAATCTATTATTTTCCGAGAGGCTGACAAATGTATAA 1022

QY 161 AsplyleucysSerTyrrArgThrAspleucysAspHlsAlaLeuHlsIleSerHlsasp 180
Db 1023 GACAAACTTTCAGTAACGCAACAGATCTTGTGACATGCCCTGCACATATCCGATGAT 1082

QY 181 Gluleu 182
Db 1083 GAGCTA 1088

RESULT 7
US-10-176-918-173
; Sequence 173, Application US/10176918
; Publication No. US2003002275A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
```

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330P1C382
; CURRENT APPLICATION NUMBER: US/10/176,918
; PRIORITY FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-918-173

Alignment Scores:
Pred. No.: 2,54e-113 Length: 1210
Score: 960.00 Matches: 182
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-10-082-502-17 (1-182) x US-10-176-918-173 (1-1210)

QY 1 MetlysglyttrpglyttrpaulaleuleuleuglyalaleuleuglythrAlatrrpAla 20
Db 543 ATGAAGGCTGGGGTGGCTGGCCCTGCTGGGGGCGCTGGGAACGGCTGGGCT 602

QY 21 ArgatgserrginaspleuHiscysglyalacysatrgalaleuValaspluencuutrrp 40
Db 603 CGAGAGACCCAGAGATCTGACCTGAGACATGAGGGCTGTGGATGAATGAACTGAATGG 662

QY 41 GluilealaglnValasprolyasrthrilegmetglsyrPhearglleaspro 60
Db 723 GATGACACCGAGTCAGTGTGGAGGTGCTTATGCCCTCAGAGGCCACCTCAGAG 782

QY 61 AspglyserginserValalgluValprotyrAlaargserglnAlaHlsleuthrglu 80
Db 843 CATGCCAAGAACTACGTACGTGTAGTGGCGCGGAATGAGAAATCCAGTAATGAGACCTA 902

QY 81 LeuleugluGluilecysAspargmetlysglyutryrglyuglnIleasproserthr 100
Db 783 CTGCTGGAGAGATATGTGACCGGATGAAGAGATGGGAAACAAATGATCTTCCACC 842

QY 101 HlsarglyAsanTyrrValargValalgluValargannglyluserSerGluDeuaspLeu 120
Db 843 CATGCCAAGAACTACGTACGTGTAGTGGCGCGGAATGAGAAATCCAGTAATGAGACCTA 902

QY 121 GlnnglyllearglleaspserraspIleserGlythrleuLysPheAlaCysgluserlle 140
Db 903 CAAGGCATCCCAATCGATCAGATATATGACGCGACCCCTCAAGTTGCGGTGAGAGCAT 962

QY 141 ValglugluTyrrGluAspGluLeuilegluphPheSerargGluAlaAspAsnVallys 160
Db 963 GTGGAGGAATACGAGAGATGAATCTATTATTTTCCGAGAGGCTGACAAATGTATAA 1022

QY 161 AsplyleucysSerTyrrArgThrAspleucysAspHlsAlaLeuHlsIleSerHlsasp 180
Db 1023 GACAAACTTTCAGTAACGCAACAGATCTTGTGACATGCCCTGCACATATCCGATGAT 1082

QY 181 Gluleu 182
Db 1083 GAGCTA 1088

RESULT 8
```

```

Db      963 GGGAGGAATACGAGCATGACACTGATTCGTTTCCGGAGAGGCTGACATGTTAA 1022
      |||||||
QY      161 ASPLYSleuCySseLysArTrhAspleuCyAspHsAlaLeuHsIleSerHisap 180
      |||||||
Db      1023 GACAAACTTTCAGTAAACGGAACAGATCTTTGTGACCATCCCTCGACATATCGCATGT 1082
      |||||||
QY      181 Glueu 182
      |||||||
Db      1083 GAGCTA 1088

RESULT 9
US-10-137-865-173
; Sequence 173, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C154
; CURRENT APPLICATION NUMBER: US/10/137,865
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-137-865-173

Alignment Scores:
Pred. NO.: 2,54e-113
Score: 960.00 Length: 1210
Percent Similarity: 100.00% Matches: 182
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-082-502-17 (1-182) x US-10-137-865-173 (1-1210)
QY      1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyAlaLeuGlyThrAlaTrpAla 20
Db      543 ATGAAAGCTGGGGGTGGCGTGGCCCTGCTTGTGGGGCCCTCTGGAAACCGCTGGGGCT 602
      |||||||
QY      21 ArgArGserGlnAspLeuHsIleCySgIyAlaCyAsArgAlaLeuValAspGluLeuGluTrp 40
      |||||||
Db      603 CGAGAGAGCCAGATCTCCACCTGTGAGCATCAGAGGCTCTGTGTGAGTGAACCTGAAATGG 662
      |||||||
QY      41 GluIleAGlnValAspProLysIlyThrIleGlnMetGlySerPheArgIleAsnPro 60
      |||||||
Db      663 GAAATTCGCCAGGTGAGCCCAAGAGACCATTCAGATGGGATCTTTCCGATCAATCCA 722
      |||||||
QY      61 AspGlySerGlnSerValValGluValProLYrAlaAarGserGluAlaHsIleuThrGlu 80
      |||||||
Db      723 GATGGCAGCCAGTCAATGTGTGGAGGTGCTTATGCCCCCTCAAGAGGCCCACTTCACAGAG 782
      |||||||
QY      81 LeuGluGlnGluIleLeuCyAspArGMetLysGluTrpGlyGluGlnIleAsnProSerThr 100

```

```
DB 783 CTGCTGGAGACATATATGTACCGGATGAAGAGATATGGGAGACAGTATCTTCACC 842
QY 101 HHSATGlyAsnTyrValArgValAlaGlyArgAsnGlyGluSerSerGluLeuAspLeu 120
DB 843 CATCGCAAGACTACTAGTGTAGTGGGCCGGAATGGAGAAATCCAGTCACTGGACCTTA 902
QY 121 GINGlyTLeaArgTLeaSerAspIleSerGlyThrLeuLeuPheAlaCysGluSerIle 140
DB 903 CAAGGATCCGATGCATCAGATCATATATAGCGCACCCCTCAAGTTGCGGTGAGACAT 962
QY 141 VALGluGluTyrGluAspGluLeuLeuIleGluPheSerArgGluAlaAspAsnValys 160
DB 963 GTGGAGAAATACAGAGATCACTGTAATTTCTTTCCGAGAGGCTGACATGTTAA 1022
QY 161 ASPlyLeuCySerLysArgThrAspLeuCyAspHisAlaLeuHisIleSerHisAsp 180
DB 1023 GACAAACTTTGAGTAAGGAGAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 1082
QY 181 GILeu 182
DB 1083 GAGCTA 1088

RESULT 10
US-10-140-474-173
; Sequence 173, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-474-173

Alignment Scores:
Pred. No.: 2,54e-113 Length: 1210
Score: 960.00 Matches: 182
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-082-502-17 (1-182) x US-10-140-474-173 (1-1210)
QY 1 MetLysGlyTyrPglYTrPleuAlaLeuLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20
DB 543 ATGAAAGGCTGGGGTGGCTGGCTGTCTTGGGGGCCCTGCTGGAAACCCCTGGGCT 602
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40
```

```
DB 603 CGAGAGAGCAGGATCTCCACTGTGAGCATGCAGGCTCTGTGGATCAACTAGATGG 662
QY 41 GILLeuAlaGluValAspProLysTyrIleGlnMetGlySerPheArgIleAsnPro 60
DB 663 GAAATGCCAGGTGAGACCCCAAGAAAGACCATTCACATGGATCTTTCCGATCAATCCA 722
QY 61 ASPGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
DB 723 GATGGAGCCAGTCAAGTGTGGTGGTGGCTTATGCCCTCAGAGGCCACCTCACAGAG 782
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100
DB 783 CTGCTGGAGACATATGTACCGGATGAAGAGATATGGGAGAACAGATTATCTTCACC 842
QY 101 HHSATGlyAsnTyrValArgValAlaGlyArgAsnGlyGluSerSerGluLeuAspLeu 120
DB 843 CATCGCAAGACTACTAGTGTAGTGGGCCGGAATGGAGAAATCCAGTCACTGGACCTTA 902
QY 121 GINGlyTLeaArgTLeaSerAspIleSerGlyThrLeuLeuPheAlaCysGluSerIle 140
DB 903 CAAGGATCCGATGCATCAGATCATATATAGCGCACCCCTCAAGTTGCGGTGAGACAT 962
QY 141 VALGluGluTyrGluAspGluLeuLeuIleGluPheSerArgGluAlaAspAsnValys 160
DB 963 GTGGAGAAATACAGAGATCACTGTAATTTCTTTCCGAGAGGCTGACATGTTAA 1022
QY 161 ASPlyLeuCySerLysArgThrAspLeuCyAspHisAlaLeuHisIleSerHisAsp 180
DB 1023 GACAAACTTTGAGTAAGGAGAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 1082
QY 181 GILeu 182
DB 1083 GAGCTA 1088

RESULT 11
US-10-142-431-173
; Sequence 173, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-431-173

Alignment Scores:
Pred. No.: 2,54e-113 Length: 1210
Score: 960.00 Matches: 182
```





APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C59  
CURRENT APPLICATION NUMBER: US/10/140,002  
CURRENT FILING DATE: 2002-05-06  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-140-002-173

Alignment Scores:  
Pred. No.: 2,546-113 Length: 1210  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-17 (1-182) x US-10-140-002-173 (1-1210)

QY 1 MetlysglytrpglytrpreunaleuLeuLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20  
DB 543 ATGAAAGCGTGGGTGGCTGGCCCTGCTTCTGGGGGCCCTCGGAACCCCTGGGCT 602  
QY 21 ArgArgSerGlnAspLeuHiscysGlyAlaCysArgAlaLeuValAspLeuLeuGluTrp 40  
DB 603 CGAGAGAGCCAGATCTCCACTGTGGAGCATGAGGCTCTGTGGATGAACTAGATGG 662  
QY 41 GllleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 663 GAAATTGCCAGGTGAGCCCAAGAACCATTCAGATGGATCTTCCGATCAATCCA 722  
QY 61 AspGlySerGlnSerValValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80  
DB 723 GATGCGACGACATGATGTGAGGTGGCTTATGCCCTCGAGAGCCACCTCACAGAG 782  
QY 81 LeuLeuGlnGluIleCysAspArgMetLysGluTyrGlyGlnGlnIleAspProSerThr 100  
DB 783 CTGCTGGAGAGATATGTACCCGATGAAGAGTATGGGAAACAGATTATCTTCCACC 842  
QY 101 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 120  
DB 843 CATCCAGAACTAGCTAGCTAGTGGGCCGGAATGGAAATCCAGTAAGTGAACCTA 902  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 903 CAAGCATCCGAAATGACACAGATATATGCCGACCTCAACTTTCCTGTAGAGCAAT 962  
QY 141 ValGlnGluTyrGlnAspGluLeuIleGluPhePheSerArgGlnAlaAspAsnValLys 160  
DB 963 GTGGAGAAATACAGAGATGAACTCATGTGAATCTTTTCCCGAGAGGCTGACATGTAAA 1022  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180  
DB 1023 GACAAACTTTGACATAGAGAAACAGATCTTTGTGACCAATGCCCTGACATATGCAATGAT 1082  
QY 181 GluLeu 182  
DB 1083 GAGCTA 1088  
RESULT 14  
US-10-142-419-173

Sequence 173, Application US/10142419  
GENERAL INFORMATION:  
US2003004945A1  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C244  
CURRENT APPLICATION NUMBER: US/10/142,419  
CURRENT FILING DATE: 2002-05-10  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-142-419-173

Alignment Scores:  
Pred. No.: 2,546-113 Length: 1210  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-17 (1-182) x US-10-142-419-173 (1-1210)

QY 1 MetlysglytrpglytrpreunaleuLeuLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20  
DB 543 ATGAAAGCGTGGGTGGCTGGCCCTGCTTCTGGGGGCCCTCGGAACCCCTGGGCT 602  
QY 21 ArgArgSerGlnAspLeuHiscysGlyAlaCysArgAlaLeuValAspLeuLeuGluTrp 40  
DB 603 CGAGAGAGCCAGATCTCCACTGTGGAGCATGAGGCTCTGTGGATGAACTAGATGG 662  
QY 41 GllleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 663 GAAATTGCCAGGTGAGCCCAAGAACCATTCAGATGGATCTTCCGATCAATCCA 722  
QY 61 AspGlySerGlnSerValValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80  
DB 723 GATGCGACGACATGATGTGAGGTGGCTTATGCCCTCGAGAGCCACCTCACAGAG 782  
QY 81 LeuLeuGlnGluIleCysAspArgMetLysGluTyrGlyGlnGlnIleAspProSerThr 100  
DB 783 CTGCTGGAGAGATATGTGACCGGATGAAGAGTATGGGAAACAGATTATCTTCCACC 842  
QY 101 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 120  
DB 843 CATCCAGAACTAGCTAGCTAGTGGGCCGGAATGGAAATCCAGTAAGTGAACCTA 902  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 903 CAAGCATCCGAAATGACACAGATATATGCCGACCTCAACTTTCCTGTAGAGCAAT 962  
QY 141 ValGlnGluTyrGlnAspGluLeuIleGluPhePheSerArgGlnAlaAspAsnValLys 160  
DB 1083 GAGCTA 1088  
RESULT 14  
US-10-142-419-173

DB 963 GTGAGCAATACGAGATGATGATTCATTGTTCTTTCCCGAGAGGCTGACAAATGTTAAA 1022  
QY 161 AspLysLeuCySerLysArgThrAspLeuCyAspHisAlaLeuHisIleSerHisAsp 180  
DB 1023 GACAAACTTGGAGTAAGGACAGATCTTTGTGACCAATGCCCTGCACATATGCAATGAT 1082  
QY 181 GluLeu 182  
DB 1083 GAGCTA 1088

## RESULT 15

US-10-123-262-173  
; Sequence 173, Application US/10123262  
; Publication No. US20030049816A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerltsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C38  
; CURRENT APPLICATION NUMBER: US/10/123,262  
; CURRENT FILING DATE: 2002-04-15  
; Prior Application removed - See file wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-123-262-173

## Alignment Scores:

Pred. No.: 2,54e-113 Length: 1210  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-10-082-502-17 (1-182) x US-10-123-262-173 (1-1210)

QY 1 MetLysGlyTTPGLYTRPLEuAlaLeuLeuGlyAlaLeuLeuGlyThraLalrPaLa 20  
DB 543 ATGAAGGCTGGGGTGGCTGGCTGCTCTGCGGGGCCCTGCTGGGAACCGCTGGGCT 602  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTTP 40  
DB 603 CGAGAGAGCCAGATCTCCACTGTGGAGCATGCGAGGCTCTGTGATGAACAGATGG 662  
QY 41 GluIleAlaGluValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 663 GAAATTTGCCAGGTGACACCCCAAGAACACCATTCACATGGATCTTCCGATCAATCCA 722  
QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 723 GATGGAGCCAGTCACTGTGTGAGGTGCTTATGCCGCTCAGAGGCCACCTCACAGAG 782  
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAsnProSerThr 100

DB 783 CTGCTGAGAGAGATATGATGACCGGATGAAGAGATATGCGGAACAGATTCCTTCACC 842  
QY 101 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGluSerSerGluLeuAspLeu 120  
DB 843 CATCGCAAGACTAGCTAGCTGTAGTGGCCGGAATGGAGAAATCCAGTGAACCTGGACCTA 902  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 903 CAAGCATCCGATCGATCGATCAGATATATAGCGGCACCCCTCAAGTTGCGTGTGAGACATT 962  
QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgLysAlaAsnValLys 160  
DB 963 GTGAGCAATACGAGATGAACTCATTCATTCTTTCCCGAGAGGCTGACAAATGTTAAA 1022  
QY 161 AspLysLeuCySerLysArgThrAspLeuCyAspHisAlaLeuHisIleSerHisAsp 180  
DB 1023 GACAAACTTGGAGTAAGGACAGATCTTTGTGACCAATGCCCTGCACATATGCAATGAT 1082  
QY 181 GluLeu 182  
DB 1083 GAGCTA 1088

Search completed: March 31, 2003, 11:59:04  
Job time : 58.0233 secs



GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:43:09 ; Search time 1166.6 Seconds

(without alignments)  
2526.643 Million cell updates/sec

Title: US-10-082-502-17

Perfect score: 960

Sequence: 1 MKGCGWLLALLGALLGTA.....LCSKRTDCLDHALHSHDEL 182

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+pn.model -DEV=xlp  
-O/cgn2.1/USPTO.spool/US10082502/runat\_24032003.135100.6599/app.query.fasta\_1.1308  
-DB=EST -OFMT=fasta -SUFFIX=est -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdt -LIST=45  
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pgn -NON-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10082502@cgn2.1.5422@runat\_24032003.135100.6599 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-NARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estinu:\*  
4: em\_estlov:\*  
5: em\_estlpl:\*  
6: em\_estro:\*  
7: em\_hic:\*  
8: em\_hic2:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inu:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_tod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	960	100.0	588	14	BM832828	BM832828 K-EST0107
2	960	100.0	591	14	BM832888	BM832888 K-EST0107
3	960	100.0	611	13	BT172899	BT172899 I696f09.y
4	960	100.0	697	12	BE344334	BE344334 602014708
5	960	100.0	719	12	BE733900	BE733900 601568370
6	960	100.0	742	12	BE824549	BE824549 602728454
7	960	100.0	745	12	BE791763	BE791763 601581586
8	960	100.0	1117	13	BM552907	BM552907 AGENCOURT
9	954	99.4	880	14	BQ421827	BQ421827 AGENCOURT
10	952	99.2	783	13	BT550436	BT550436 603192905
11	947	98.6	868	14	BQ421069	BQ421069 AGENCOURT
12	945	98.4	791	12	BT700443	BT700443 602688633
13	939	97.8	773	13	BT550162	BT550162 603193045
14	938	97.7	754	11	AK013568	AK013568 Mus muscu
15	938	97.7	772	11	AK007914	AK007914 Mus muscu
16	938	97.7	808	13	BT408525	BT408525 602964995
17	938	97.7	843	13	BT412539	BT412539 602990995
18	938	97.7	853	11	AK013014	AK013014 Mus muscu
19	938	97.7	1255	11	AK019927	AK019927 Mus muscu
20	934	97.3	724	14	BQ210416	BQ210416 UT-R-DY1-
21	934	97.3	818	12	BE870315	BE870315 601447464
22	934	97.3	875	13	BT763950	BT763950 603049829
23	933	97.2	887	13	BT762759	BT762759 603048045
24	927	96.6	673	14	BQ668530	BQ668530 AGENCOURT
25	927	96.6	720	13	BT909759	BT909759 603070841
26	926	96.5	802	12	BF608664	BF608664 MYL_00160
27	925	96.4	891	13	BT408381	BT408381 602964013
28	920	95.8	783	13	BT763784	BT763784 603049729
29	920	95.8	914	12	BE799495	BE799495 601589289
30	911	94.9	765	12	BF312281	BF312281 601898804
31	906	94.4	856	12	BF965088	BF965088 602268856
32	901	93.9	665	10	AW919569	AW919569 EST350873
33	900	93.8	856	13	BT454180	BT454180 603170602
34	899	93.6	791	12	BT456021	BT456021 603170466
35	898	93.5	751	12	BE742698	BE742698 601574753
36	892	92.9	709	13	BT411532	BT411532 602965126
37	887	92.4	797	13	BT762927	BT762927 603047987
38	883	92.0	885	14	BQ920767	BQ920767 AGENCOURT
39	883	92.0	963	14	BQ929020	BQ929020 AGENCOURT
40	880	91.7	908	14	BQ877139	BQ877139 AGENCOURT
41	880	91.4	1019	12	BG024176	BG024176 602303060
42	877	91.4	776	13	BT522523	BT522523 603175203
43	874	91.0	699	9	AL548119	AL548119 AT548119
44	869	90.5	746	13	BT551339	BT551339 603193144
45	867	90.3	930	14	BQ437660	BQ437660 AGENCOURT

## ALIGNMENTS

RESULT 1  
LOCUS BM832828  
DEFINITION K-EST0107375 S55NU484s1 Homo sapiens cDNA clone S55NU484s1-11-B03  
5', mRNA sequence.  
ACCESSION BM832828  
VERSION BM832828.1 GI:19189237  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 588)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S.

TITLE  
JOURNAL  
COMMENT

FEATURES

Kim, Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsungemail.kr@kribb.re.kr  
Plate: 11 row: B column: 03  
High quality sequence stop: 588.  
Location/Qualifiers

BASE COUNT

149 a 140 c 176 g 123 t

Alignment Scores:

Pred. NO.: 1.75e-111 Length: 588  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-10-082-502-17 (1-182) x BM832828 (1-588)

QY 1 MettysglttptglttptleuAlaleuLeuGlyAlaleuLeuGlyThrAlaTtrpAla 20  
Db 2 ATGAAGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 61  
QY 21 ArgatgserGlnAspLeuHnscysGlyAlaCysArgAlaLeuValAspGluLeuGluTtrp 40  
Db 62 CGGAGACCCAGGATCTCCACTGTGGACATGCAAGGCTGTGGTGAATGAATGAATGG 121  
QY 41 GluilealaglnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
Db 122 GAAATTCGCCAGGTGGACCCCAAGAACCAATTCGATGGGATCTTTCCGGATCAATCA 181

QY 61 AspGlySerGlnSerValValGluValProTyrrAlaArgSerGluAlaHsLeuThrglu 80  
Db 182 GATGGACGACCACTGATGGTGGAGGCTGTATGCCCGCTCAGAGGCCACCTCACAGAG 241  
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrrGlyGluGlnIleAspProSerThr 100  
Db 242 CTGCTGGAGAGATATATGTCACCGAGTGAAGAGTATGGGGAAACATATCATCTTCACAC 301  
QY 101 HsArgLysAsnTyrrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 120  
Db 302 CATCCAGAACATACGATACGATACGATACGATACGATACGATACGATACGATACGAT 361  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 362 CAAGCATCCGAAATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 421  
QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
Db 422 GTGAGGAATACGAGATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 481  
QY 161 AspLysLeuCySerLysArgThrAspLeuCyAspHsAlaLeuHsIleSerHsAsp 180  
Db 482 GACAAACTTTCAGATGACGACGACGACGACGACGACGACGACGACGACGACGACGAT 541  
QY 181 GluLeu 182  
Db 542 GAGCTA 547

RESULT 2  
BM832888  
LOCUS  
DEFINITION  
K-EST0107445 S5SNU484s1 Homo sapiens cDNA clone S5SNU484s1-11-H01  
5', mRNA sequence.  
ACCESSION  
BM832888  
VERSION  
BM832888.1 GI:19189297  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 591)  
Kim, Y.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
Oh, K.U., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
Kim, Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsungemail.kr@kribb.re.kr  
Plate: 11 row: H column: 01  
High quality sequence stop: 591.  
Location/Qualifiers

FEATURES

source

1. 591  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="S5SNU484s1-11-H01"  
/clone\_lib="S5SNU484s1"  
/sex="M"  
/tissue\_type="Stomach"  
/cell\_type="Epithelial"  
/lab\_host="Top10F"  
/note="Organ: Stomach; Vector: pTZ189P1; Site\_1: EcoRI;  
Site\_2: NotI; The poly (A)+ RNA was decapped with tobacco  
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
including EcoRI site by treatment of T4 RNA ligase. The  
first strand cDNA was synthesized from oligo dt-selected  
mRNA by priming with dt-tailed vector. The dt-tailed  
vector was adjusted to have about 60nt. The cDNA vector

was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(dn)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by the transformation of the remaining DNA into competent cells E. coli Top10F' with electroporation method."

BASE COUNT 150 a 142 c 176 g 123 t  
ORIGIN

## Alignment Scores:

Pred. No.: 1,76e-111 Length: 591  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-10-082-502-17 (1-182) x BM632888 (1-591)

QY 1 MetLysGlyTyrPglYtrPLeuAlaLeuLeuGlyAlaLeuGlyThrAlaTTPala 20  
|||||  
DB 2 ATGAAAGCGTGGGGTGGCGCCCTGCTCTGGGGCCCTGGGAACCGCTGGGCT 61  
21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTTP 40  
|||||  
DB 62 CGAGAGAGCGCAGATCTCCACTGTGAGCATGAGGCTTGATGATGATGATGATGATG 121  
41 GluTLeuAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
|||||  
DB 122 GAATATGCCAGGTGAGACCCCAAGAACCATTCAGATGGATCTTCGGGATCAATCCA 181  
61 AspGlySerGlnSerValValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80  
|||||  
DB 182 GATGGAGCCAGTCACTGTGTGGAGTGCCTTATGCCCGCCAGAGGCCACTCAGAGAG 241  
81 LeuLeuGluGluTLeuCysAspArgMetLysGluTyrGluGluGlnIleAspProSerThr 100  
|||||  
DB 242 CTCCTGGAGAGATATGTGTGACCGGATGAGAGATATGGGAGACAGATGATCCCTCCACC 301  
101 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGluSerSerGluLeuAspLeu 120  
|||||  
DB 302 CATCCGAAAGAACTACGTACGTAGTGGGCCGGAATGAGAAATCCAGTGAATCGAACCTA 361  
121 GlnGlyIleArgGluLeuAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
|||||  
DB 362 CAAGGATCCGAAATGCACTGATATATAGCGGACCCCTCAAGTTCCGTGTGAGACATTT 421  
141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
|||||  
DB 422 GTGGAGGAATACGAGATGATGATCAATTTTTCGAGAGAGCTGACAAATGTTTAA 481  
161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180  
|||||  
DB 482 GACAAACTTTCAGTAAAGGAGACAGATCTTTGTGACCATGCCCTGCACATATGCATGAT 541  
181 GluLeu 182  
542 GAGCTA 547

RESULT 3  
B1712899

LOCUS B1712899 611 bp mRNA linear EST 11-MAR-2002  
DEFINITION id98f09.y1 HR65 1slet Homo sapiens cDNA clone IMAGE:5085832 5'

similar to TR:Q9Y2B0 Q9Y2B0 TYPE II MEMBRANE PROTEIN. ; mRNA

sequence:

ACCESSION B1712899

VERSION B1712899.1

KEYWORDS GT:15688594

SOURCE EST.

ORGANISM human.

REFERENCE Mammalia: Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS 1 (bases 1 to 611)

Mellon/D., Brown/J., Kenty/G., Pernutt/A., Lee/C., Kaestner/R.,

Lemishka/I., Scaevole/M., Brestelli/J., Gradwohl/G., Clifton/S.,

Hillier/L., Maria/M., Pape/D., Wylie/T., Martin/J., Blistain/A.,

Schmitt/A., Theising/B., Rifter/E., Ronko/I., Bennett/J., Cardenas

M., Gibbons/M., McCann/R., Cole/R., Tsagaris/Will,I., Williams,T.,

Jackson,X. and Bowers,Y.

Endocrine Pancreas Consortium

Unpublished (2000)

CONTACT: Douglas Mellon, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept. of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmellon@biohp.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:

Washington University Genome Sequencing Center For information on

obtaining a clone please contact: Dr. Hiroshi Inoue

(hinoue@im.wustl.edu)

Seq primer: -40RP from Gibco

High quality sequence stop: 463.

Location/Qualifiers

1..611

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5085832"

/clone\_lib="HR65 1slet"

/tissue\_type="Purified pancreatic 1slet"

/lab\_host="DH10B"

/note="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1:

NotI; Site\_2: XhoI; cDNA made by oligo-dT priming.

Size-selected on agarose gel. Average insert size

XhoI site was destroyed after directional cloning.

Amplified once. Contact information: Hiroshi Inoue, MD,

Metabolism Div. (Alan Permutt Lab), Washington University

School of Medicine, Box 8127, 660 South Euclid Ave., St.

Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:

314-362-1916, Fax: 314-747-2692."

BASE COUNT 157 a 143 c 184 g 127 t

ORIGIN

Alignment Scores:

Pred. No.: 1.85e-111 Length: 611

Score: 960.00 Matches: 182

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 13 Gaps: 0

US-10-082-502-17 (1-182) x B1712899 (1-611)

QY 1 MetLysGlyTyrPglYtrPLeuAlaLeuLeuGlyAlaLeuGlyThrAlaTTPala 20  
|||||  
DB 30 ATGAAAGCGTGGGGTGGCGCCCTGCTCTGGGGCCCTGGGAACCGCTGGGCT 89  
21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTTP 40  
|||||  
DB 90 CGAGAGAGCGCAGATCTCCACTGTGAGCATGAGGCTTGATGATGATGATGATGATG 149

QY 41 GluTLeuAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60

|||||  
DB 150 GAAATTGCCAGGTGAGCCCAAGAGACCATTCAGATGGATCTTCCGGATCAATCCA 209  
OY 61 ASpglySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 210 GATGGCAGCCAGTCACTAGTGTGAGAGTGCCTTATGCCCGCTCAGAGCCCACTCAGAG 269  
OY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGluGluGlnIleAspProSerThr 100  
DB 270 CTGCTGGAGAGATATGTGACCGCATGAAGAGATGTGGGAACAGATGATCTTCCACC 329  
OY 101 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 120  
DB 330 CATCCCAAGAACTACGTACGTGTAGTGGCCGGAATGAGAAATCCAGAACTGGACCTA 389  
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 390 CAAGGCATCCGAATGAGTCAAGATATTAGCGGCACCCCTCAAGTTGCCGTGAGAGCAT 449  
OY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
DB 450 GTGGAGGATACGAGGATCAACTCATTAATCTTTCCGAGAGCGCTGACAAATGTAAA 509  
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspPheAlaLeuHisIleSerHisAsp 180  
DB 510 GACAAACTTTCAGTACGAGAACAGATCTTGTGACACATGCCCTGCACATATGCATGAT 569  
OY 181 GluLeu 182  
DB 570 GAGCTA 575  
RESULT 4  
BF344334 697 bp mRNA linear EST 22-NOV-2000  
LOCUS 602014708F1 NCI\_CGAP\_Brn64 Homo sapiens cDNA clone IMAGE:4150512  
DEFINITION 5', mRNA sequence.  
ACCESSION BF344334  
KEYWORDS BF344334.1 GI:11291554  
SOURCE EST.  
ORGANISM Homo sapiens  
human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 697)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LNLN)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
plate: LLM9413 row: P column: 01  
High quality sequence stop: 692.  
Location/Qualifiers  
1. .697  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4150512"  
/clone\_id="NCI\_CGAP\_Brn64"  
/tissue\_type="globlastoma with EGFR amplification"  
/lab\_host="DH10B (TI phage-resistant)"  
/note="Organ: brain; Vector: PCMV-SPORT6; Site\_1: NCI;  
Site\_2: Salt; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.57 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 173 a 167 c 220 g 137 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,25e-111 Length: 697  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0  
US-10-082-502-17 (1-182) x BF344334 (1-697)  
OY 1 MetLysGlyTyrPglTyrPLeuAlaLeuLeuLeuGluAlaLeuLeuGlyThrAlaTrrPala 20  
DB 92 ATGAAAGCTGGGGTGGCTGGCCCTGCTTGTGGGGCCCTGGTGGAAACCGCTGGCT 151  
OY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrrP 40  
DB 132 CGGAGAGCCAGAGATCTCACAGTGTGAGACATCCAGCGCTCTGGTGAAGAACATGAG 211  
OY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 212 GAAATTGCCAGGTGAGCCCAAGAGACCATTCAGATGGATGATCTTCCGGATCAATCCA 271  
OY 61 ASpglySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 272 GATGGCAGCCAGTCACTAGTGTGAGAGTGCCTTATGCCCGCTCAGAGCCCACTCAGAG 331  
OY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGluGluGlnIleAspProSerThr 100  
DB 332 CTGCTGGAGAGATATGTGACCGCATGAAGAGATGTGGGAACAGATGATCTTCCACC 391  
OY 101 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 120  
DB 392 CATCCCAAGAACTACGTACGTGTAGTGGCCGGAATGAGAAATCCAGTGAACCTGGACCTA 451  
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 452 CAAGGCATCCGAATGAGTCAAGATATTAGCGGCACCCCTCAAGTTGCCGTGAGAGCAT 511  
OY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
DB 512 GTGGAGGATACGAGGATCAACTCATTAATCTTTCCGAGAGCGTACAAATGTAAA 571  
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspPheAlaLeuHisIleSerHisAsp 180  
DB 572 GACAAACTTTCAGTACGAGAACAGATCTTGTGACACATGCCCTGCACATATGCATGAT 631  
OY 181 GluLeu 182  
DB 632 GAGCTA 637  
RESULT 5  
BE733900 719 bp mRNA linear EST 15-SEP-2000  
LOCUS 601568370F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:3843209 5',  
DEFINITION mRNA sequence.  
ACCESSION BE733900  
KEYWORDS BE733900.1 GI:10147892  
SOURCE EST.  
ORGANISM Homo sapiens  
human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 719)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin laboratory  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be



found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)  
Plate: LLCM537 row: k column: 18  
High quality sequence stop: 710.  
Location/Qualifiers

## FEATURES

source

1. 719  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="3843209"  
/clone\_lib="NIH\_MGC\_21"  
/issue\_type="choriocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: placenta; Vector: pORF7; Site:1: XhoI;  
Site:2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAAGAG(g). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 190 a 164 c 214 g 151 t  
ORIGIN

## Alignment Scores:

Pred. No.:	2,356-111	Length:	719
Score:	960.00	Matches:	182
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
	12	Gaps:	0

US-10-082-502-17 (1-182) x BE733900 (1-719)

```
QY 1 MetlySGlyTTPpGlyTTPLeuAlaLeuLeuLeuGlyAlaLeuLeuGlyThrAlaTTPAla 20
Db ATAAAGCGTGGGGTGGCGCCCTCTCTGCGGGCCCTGCGGAGACCGCGTGGGCT 107
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTTP 40
Db 108 CGAGAGAGCCAGATCTCCCTGCGAGCATGCCAGGCTGTGTGATGAGTGAATGG 167
QY 41 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60
Db 168 GAATATGCCAGGTGGACCCCAAGAGACCATTCAGATGGAGATCTTCCGAGTCAATCCA 227
QY 61 AspGlySerGlnSerValValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
Db 228 GATGGAGCCAGTCAGTGGTGGAGGTGCTTATGCCGCTCAGAGGCCACCTCACAGAG 287
QY 81 LeuLeuGluGluIleCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100
Db 288 CTGCTGGAGAGATATCTGACCCGATGACAGAGATATGGGACAGATTTGCTTCCACC 347
QY 101 HisArgLysAsnTyrValAlaArgValValGlyArgAsnGlyLysSerGluLeuAspLeu 120
Db 348 CATCGGAGAACTACGACGTAGTGGGCCGGAATGGAATCGAATCGAATCGAGACTTA 407
QY 121 GlnGlyIleLeuArgLysPheAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 408 CAAGGATCCGAAATCGACTCAGATATTAGCGGACCCCTCAAGTTGGCTGTGAGAGACTT 467
QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160
Db 468 GTGGAGAGATACAGAGATGACATCATTTATTTTCCCGAAGAGGCTGCAATGTATAA 527
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180
Db 528 GCAAAACTTTGACAGTAAGCAGATCTTTGTGACATGCTCCGACATATGCGATGAT 587
QY 181 GluLeu 182
Db 588 GAGCTA 593
RESULT 6
```

BG824549

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

602728454F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4868100 5',  
BG824549  
EST. 742 bp mRNA linear EST 22-MAY-2001

Human.  
Homo sapiens  
Chordata: Craniata: Vertebrata: Euteleostomi:  
Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.  
1 (bases 1 to 742)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: NIH Intramural Sequencing Center  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM1737 row: k column: 13  
High quality sequence stop: 740.  
Location/Qualifiers

1. 742  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="4868100"  
/clone\_lib="NIH\_MGC\_15"  
/issue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: colon; Vector: pORF7; Site:1: XhoI; Site:2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCAAGAG(g). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 180 a 184 c 233 g 144 t  
ORIGIN

## Alignment Scores:

Pred. No.:	2,476-111	Length:	742
Score:	960.00	Matches:	182
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
	12	Gaps:	0

US-10-082-502-17 (1-182) x BG824549 (1-742)

```
QY 1 MetlySGlyTTPpGlyTTPLeuAlaLeuLeuLeuGlyAlaLeuLeuGlyThrAlaTTPAla 20
Db 141 ATAAAGCGTGGGGTGGCGCCCTCTCTGCGGGCCCTGCGGAGACCGCGTGGGCT 200
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTTP 40
Db 201 CGAGAGAGCCAGATCTCCCTGCGAGCATGCCAGGCTGTGTGATGAGTGAATGG 260
QY 41 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60
Db 261 GAATATGCCAGGTGGACCCCAAGAGACCATTCAGATGGAGATCTTCCGAGTCAATCCA 320
QY 61 AspGlySerGlnSerValValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
Db 321 GATGGAGCCAGTCAGTGGTGGAGGTGCTTATGCCGCTCAGAGGCCACCTCACAGAG 380
QY 81 LeuLeuGluGluIleCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100
Db 81 LeuLeuGluGluIleCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100
```

DB 381 CTGCTGAGAGATATGTGACCGGATGAAGAGATGCGGAACAGATTGATCCTTCACAC 440  
QY 101 HIsAgLySaNTyrValArgValAlGlyArgAsnGlyGluSerSerGluLeuAspLeu 120  
DB 441 CATCCCAAGAACTACGTACGTGTACTGGCCCGAATGGAGATCCAGAACTGACCTTA 500  
QY 121 GInGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 501 CAGGCAATCCGAAATGACTCAGATATATAGCGGACCCCTCAAGTTGGCTGTAGAGCAT 560  
QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
DB 561 GGGAGGAATACGAGAGATGAAGTCAATTCATCTTTCCCGAGAGCGTACAAATGTTAA 620  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspPheAlaLeuHisIleSerHisasp 180  
DB 621 GACAAACTTTCAGTACGAGAACAGATCTTGTGACCATGCGCTGCACATATGCGATGAT 680  
QY 181 GluLeu 182  
DB 681 GAGCTA 686

RESULT 7  
BE791763 745 bp mRNA linear EST 20-SEP-2000  
LOCUS BE791763  
DEFINITION BE791763 NIH\_MGC\_7 Homo sapiens CDNA clone IMAGE:3936141 5',  
mRNA sequence.  
ACCESSION BE791763  
VERSION BE791763.1 GI:10212961  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 745)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: DCTD/DTP  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at: [image.llnl.gov](http://image.llnl.gov)  
Plate: LCM779 row: k column: 22  
High quality sequence stop: 739.  
Location/Qualifiers  
1..745  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3936141"  
/clone\_1lb="NIH\_MGC\_7"  
/tissue-type="small cell carcinoma"  
/cell\_line="MGC3"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pOTB7; site\_1: XhoI; site\_2:  
EcoRI; CDNA made by oligo-dt priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGGACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the Laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-CDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 182 a 180 c 237 g 146 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,48e-111 Length: 745  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0  
US-10-082-502-17 (1-182) x BE791763 (1-745)  
QY 1 MetLysGlyTTPGlyTyrPheLeuAlaLeuLeuGluAlaLeuGluGlyThrAlaSerPhe 20  
DB 138 ATGAAAGGCTGGGGTGGCTGGCCCTGCTTCTGGGGGCCCTGCTGGAAACCGCTGGGCT 197  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysAlaGlyAlaLeuValAspGluLeuGlyTyr 40  
DB 198 CGGAGAGACCGAGATCTCCACTGTGAGACATGACAGGCTGTGGTGAATGAACTGAATGCG 257  
QY 41 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgTLeuAsnPro 60  
DB 258 GAATATGCCAGGTGGAGCCCAAGAAACCAATTCAGATGGGATCTTTCGGGATCAATCCA 317  
QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisIleuThrGlu 80  
DB 318 GATGGACCCAGCTCAGTGTGGAGGTGCTTATGCTCCCTCAGAGGCCCTCAGACAG 377  
QY 81 LeuLeuGluGluIleCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100  
DB 378 CTGCTGAGAGAGATATGTGACCGGATGAAGAGATGTGGGAAACAGATTGATCTTCACAC 437  
QY 101 HIsAgLySaNTyrValArgValAlGlyArgAsnGlyGluSerSerGluLeuAspLeu 120  
DB 438 CATCCCAAGAACTACGTACGTGTACTGGCCCGAATGGAGATCCAGAACTGACCTTA 497  
QY 121 GInGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 498 CAGGCAATCCGAAATGACTCAGATATATAGCGGACCCCTCAAGTTGGCTGTAGAGCAT 557  
QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
DB 558 GGGAGGAATACGAGAGATGAAGTCAATTCATCTTTCCCGAGAGCGTACAAATGTTAA 617  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspPheAlaLeuHisIleSerHisasp 180  
DB 618 GACAAACTTTCAGTACGAGAACAGATCTTGTGACCATGCGCTGCACATATGCGATGAT 677  
QY 181 GluLeu 182  
DB 678 GAGCTA 683

RESULT 8  
BM552907 1117 bp mRNA linear EST 20-FEB-2002  
LOCUS BM552907  
DEFINITION AGNCOURT\_6572552 NIH\_MGC\_41 Homo sapiens CDNA clone IMAGE:5466942  
5', mRNA sequence.  
ACCESSION BM552907  
VERSION BM552907.1 GI:18791172  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1117)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: DCTD/DTP  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.llnl.gov>  
Plate: LCM1968 row: c column: 07  
High quality sequence stop: 709.  
Location/Qualifiers

FEATURES

```

source
1. 1117
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="IMAGE:5466942"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: skin; Vector: pORF7; Site: 1: XhoI; Site: 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT      299 a      258 c      320 g      239 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:      4.53e-111      Length:      1117
Score:          960.00      Matches:      182
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             13      Gaps:        0

US-10-082-502-17 (1-182) x BM552907 (1-1117)

QY 1 MetlysglyttrpglyttrpleualaaleuLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20
Db 148 ATGAAGGCGTGGGTTGGCTGGCCCTGCTGGGGGCGCTGGGAGACCGCTGGGCT 207
QY 21 ArgArSerGlnAAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40
Db 208 CGGAGGAGCCGAGATCTCCACTGTGGAGCAGAGGCGCTGGTGATGAAGTGAATGG 267
QY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 60
Db 268 GAAATTTGCCCGAGTGGACCCCAAGAACCATTCAGATGATCTTCCGGATCAATCCA 327
QY 61 AspglySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
Db 328 GATGGCAGCCAGTCACTGTGGAGTGCCTTATCCCGTCAGAGGCCACCTCAGAGAG 387
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100
Db 388 TGTCTGAGAGAGATATGTGACCGGATGAGAGAGTATGGGAAACGATTCCTCCACC 447
QY 101 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyGluSerSerGluLeuAspLeu 120
Db 448 CATGGCAAGAACTACGTACGTAGTGGGCGGAGATGAGAAATCCAGTGAACCTGACCTA 507
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 508 CAAAGGCAATCCGAATCGACTGAGATTAAGCGGACCTCAAGTTTGGCTGTGAGAGCAT 567
QY 141 ValGluGluIleTyrGlnAspGluLeuIleGluPhePheSerArgGluIleAspAsnValLys 160
Db 568 GTGAGAGAAATACGAGATGAGATCATTCATTTTCCCGAAGGGGTGCAATGTGTA 627
QY 161 AspglySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
Db 628 GAAATTTGCCCGAGTGGACCCCAAGAACCATTCAGATGATCTTCCGGATCAATCCA 281
QY 181 GluLeu 182
Db 688 GAGCTA 693

RESULT 9
LOCUS      BQ421827      880 bp      mRNA      linear      EST 23-MAY-2002
DEFINITION AGENCOURT.7801623 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5050568
5', mRNA sequence.

ACCESSION      BQ421827
VERSION        BQ421827.1      GI:21117142
KEYWORDS
SOURCE
ORGANISM      human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 880)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: rgs@psh.rii.nih.gov
Tissue Procurement: ATCC/DCTD/DMR
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM3303 row: a column: 01
High quality sequence stop: 580.
Location/Qualifiers
1. 880
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="IMAGE:5050568"
/clone_1ib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: skin; Vector: pCMV-Sport6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."
Technology: "

BASE COUNT      238 a      208 c      256 g      176 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:      1.84e-110      Length:      880
Score:          954.00      Matches:      181
Percent Similarity: 99.45%      Conservative: 0
Best Local Similarity: 99.45%      Mismatches: 1
Query Match:    99.38%      Indels:      0
DB:             14      Gaps:        0

US-10-082-502-17 (1-182) x BQ421827 (1-880)

QY 1 MetlysglyttrpglyttrpleualaaleuLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20
Db 102 ATGAAGGCGTGGGTTGGCTGGCCCTGCTGGGGGCGCTGGGAGACCGCTGGGCT 161
QY 21 ArgArSerGlnAAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40
Db 162 CGGAGGAGCCGAGATCTCCACTGTGGAGCAGAGGCGCTGGTGATGAAGTGAATGG 221
QY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 60
Db 222 GAAATTTGCCCGAGTGGACCCCAAGAACCATTCAGATGATCTTCCGGATCAATCCA 281
QY 61 AspglySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
Db 282 GATGGCAGCCAGTCACTGTGGAGTGCCTTATCCCGTCAGAGGCCACCTCAGAGAG 341
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100
Db 342 GTGCTGAGAGAGATATGTGACCGGATGAGAGAGTATGGGAAACGATTCCTCCACC 401
QY 101 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyGluSerSerGluLeuAspLeu 120
Db 402 CATCCCAAGAACTACGTACGTAGTGGGCGGATGAGAAATCCAGTGAACCTGACCTA 461
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140

```

Db 462 CAAGCATCCGATCGACTGACATATTTAGCGCACCCCTCAAGTTTCCGTGTAGACATT 521

Qy 141 ValgluIuTyrgLuAspIuLeuIleGluphPheSerArgLuAlaAspAsnVallys 160

Db 522 GTGGAGGATACGAGGATGACTCATTTGAACTCTTTCCGAGAGCTGACATGTANA 581

Qy 161 AspIysLeuCySerIysArgThrAspLeuCyAspHisAlaLeuHisIleSerHisasp 180

Db 582 GACAAACTTTCAGTAAGGACAGATCTTTGTGACCATGCGCTGCACATATCGCATGAT 641

Qy 181 GluLeu 182

Db 642 GAGCTA 647

RESULT 10

BI550436 783 bp mRNA linear EST 05-SEP-2001

LOCUS 60319205F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:5264328 5',

DEFINITION mRNA sequence.

ACCESSION BI550436

VERSION BI550436.1 GI:15437748

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 783)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cgaabs-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki Yoshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://limage.llnl.gov Plate: LLM11666 row: 1 column: 01 High quality sequence stop: 730.

FEATURES

source

1..783

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_image="5264328"

/clone\_lib="NIH\_MGC\_95"

/tissue\_type="hippocampus"

/lab\_host="DH10B"

/note="Organ: Brain; Vector: pBluescriptPR (modified pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcagag); Oligo-dt primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.5 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 209 a 189 c 225 g 160 t

ORIGIN

Alignment Scores:

Pred. No.: 2.77e-110 Length: 783

Score: 952.00 Matches: 181

Percent Similarity: 99.45% Conservative: 0

Best Local Similarity: 99.45% Mismatches: 1

Query Match: 99.17% Indels: 0

Db: 13 Gaps: 0

US-10-082-502-17 (1-182) x BI550436 (1-783)

Qy 1 MetIysGILYTPGILYTPLeuAlaLeuLeuGILYAlaLeuLeuGILYThraIaTTPala 20

Db 84 ATGAAGGCTGGGGCTTGGCTGGCCCTGCTTCTGGGGGCCCTGCGAACCCTCGGGCT 143

Qy 21 ArgAspSerGluAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuIuTP 40

Db 144 CGGAGGAGCCAGGATCTCCACTGTGGAGCATCGAGGCTCTGTGGATGAACTACAAATGG 203

Qy 41 GluIleAlaGluValAspProIysArgThrIleGlnMetGlySerPheArgIleAsnPro 60

Db 204 GAAATTTGCCAGAGTGGAGCCCAAGAAAGACCATTCAGATGGATCTTCCGGATCAATCCA 263

Qy 61 AspGlySerGlnSerValValGluValProIyAlaIaArgSerGluAlaHisLeuThrGlu 80

Db 264 GATGGCAGCCAGCTCATGTGGTGGAGGTGCTTATGCCCCGCTAAGAGGCCACCTCACAGG 323

Qy 81 LeuLeuGluGluIleCysAspArgMetIysGluTyrgIuGluGlnIleAspProSerThr 100

Db 324 CTGCTGGAGAGATATGTGCACCGGATGAAAGAGCTATGGGAAACAAATTCCTTCACCC 383

Qy 101 HisArgIysAsnTyrgValArgValValGlyArgAsnGlyIuSerSerGluLeuAspLeu 120

Db 384 CATGCCAAGAACTACGTACGTAGTGGCCGGAATGGAGATCCAGTGAACCTGAGACCTA 443

Qy 121 GInGlyIleArgIleAspSerAspIleSerGlyThrLeuIysPheAlaCysGluSerIle 140

Db 444 CAAGCATCCGATCGACTGACATATTTAGCGCACCCCTCAAGTTTCCGTGTAGACATT 503

Qy 141 ValgluIuTyrgLuAspIuLeuIleGluphPheSerArgLuAlaAspAsnVallys 160

Db 504 GTGGAGGATACGAGGATGACTCATTTGAACTCTTTCCGAGAGCTGACATGTANA 563

Qy 161 AspIysLeuCySerIysArgThrAspLeuCyAspHisAlaLeuHisIleSerHisasp 180

Db 564 GACAAACTTTCAGTAAGGACAGATCTTTGTGACCATGCGCTGCACATATCGCATGAT 623

Qy 181 GluLeu 182

Db 624 GAGCTA 629

RESULT 11

BQ421069 868 bp mRNA linear EST 23-MAY-2002

LOCUS BQ421069

DEFINITION AGENCOURT.7911956 NIH\_MGC\_68 Homo sapiens cDNA clone IMAGE:6009949

ACCESSION BQ421069

VERSION BQ421069.1 GI:21116384

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 868)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cgaabs-remail.nih.gov Tissue Procurement: DCTD/DTF/Gazdar cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://limage.llnl.gov Plate: LLM13197 row: d column: 14 High quality sequence stop: 542.

FEATURES

source

1..868

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_image="6009949"

/clone\_lib="NIH\_MGC\_68"

/tissue\_type="large cell carcinoma"



Db 619 GTGAGAGATAGAGAGATCAATGATCTTTTCCCGAGAGCGTACATGTAA 678  
Qy 161 AspyLseUcySerLysArgThrAspLeuCyAspHisAlaLeuHisSerHisasp 180  
Db 679 GACAAACCTTGAGTAAGCAGACAGATCTTGTGACACATGCCCTGCACATTCGATGAT 738  
Qy 181 GAGCTA 182  
Db 739 GAGCTA 744  
RESULT 13  
BI550162 773 bp mRNA linear EST 05-SEP-2001  
LOCUS BI550162  
DEFINITION 603193045F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:5264339 5',  
mRNA sequence.  
ACCESSION BI550162  
VERSION BI550162.1 GI:15437474  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 773)  
NIH-MGC http://mhc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
AUTHORS Contact: Robert Strausberg, Ph.D.  
EMAIL: cga@bbs-rtmail.nih.gov  
Tissue Procurement: Miklos Paikovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihaki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: LLM11666 row: 1 column: 12  
High quality sequence stop: 765.  
FEATURES  
Location/Qualifies  
source 1..773  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5264339"  
/clone\_1lb="NIH\_MGC\_95"  
/cissue\_type="hippocampus"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pBluescriptR (modified  
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag  
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
size-selected for average insert size 2.5 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."  
BASE COUNT 204 a 178 c 226 g 163 t 2 others  
ORIGIN  
Alignment Scores:  
Pred. NO.: 1.22e-108 Length: 773  
Score: 939.00 Matches: 181  
Percent Similarity: 98.91% Conservative: 0  
Best Local Similarity: 98.91% Mismatches: 1  
Query Match: 97.81% Indels: 1  
Gaps: 0  
US-10-082-502-17 (1-182) x BI550162 (1-773)  
Qy 1 MetLysGlyTyrPGLYTrpLeuAlaLeuLeuGlyAlaLeuGlyThrAlaTrpAla 20  
Db 84 ATGAAGAGCTGGGGTGGCTGCTGCTGGGGGGCCCTGGAACCGCTGGGCT 143  
Qy 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrrp 40

Db 144 CGAGAGCGCAGAGATCTCCAGCTGAGCATGACAGGCTGTGTGATGAACTACAAATGG 203  
Qy 41 GUILleAGlValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
Db 204 GAAATTGCCAGCTGAGCACCACCAAGAACCATTCACATGGATCTTTCCGGATCATCA 263  
Qy 61 AspglySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
Db 264 GATGCAGCAGCACTAGTGTGTGAGGTGCTTATGCCCGCTAGAGGCCACCTCACAGAG 323  
Qy 81 LeuLeuGluGluIleCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100  
Db 324 GTGCTGAGAGATATGTGATCCGATGAAGAGTGTGGGAAACAATGATCTCTCCACAC 383  
Qy 101 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerGluLeuAspLeu 120  
Db 384 CATCCCAAGAACTACGCTAGCTAGTGTAGTGGCCGGAATGGAAATCCAGTGAACCTGA 443  
Qy 121 GInGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 444 CAAGCATCCGAATGAGTCAAGATATAGCGCACCTCAAGTTTGCCTGAGAGCAT 503  
Qy 141 ValGluGluTyrGluAspGluLeuLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
Db 504 GTGAGGAATACGAGAGATGAACTCATGTGAATCTTTTCCCGAGAGCGTGCACAAATGTTAA 563  
Qy 161 AspyLseUcySerLysArgThrAspLeu-CysAspHisAlaLeuHisIleSerHisasp 180  
Db 564 GACAAACCTTGAGTAAGCAGACAGATCTTGTGACACATGCCCTGCACATATCCATCA 623  
Qy 180 pGluLeu 182  
Db 624 TGAGCTA 630  
RESULT 14  
LOCUS AK013568  
DEFINITION Mus musculus adult male hippocampus cDNA, RIKEN full-length  
inserted library, clone:2900022B12:transmembrane protein 4, full  
insert sequence.  
ACCESSION AK013568.1 GI:12850979  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (strain:C57BL/6J) adult male hippocampus cDNA to mRNA,  
clone:2900022B12.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PubMed 10349636  
REFERENCE 2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PubMed 11042159  
REFERENCE 3  
Shibata, K., Itoh, M., Mizawa, K., Nagaoke, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, U., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,  
Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Iwata, M., Ohara, E., Watabiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, T., Kita, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format



RESULT	15
LOCUS	AK007914
DEFINITION	Mus musculus 10 day old male pancreas cDNA, RIKEN full-length insert sequence.
ACCESSION	AK007914
VERSION	AK007914.1 GI:12841773
SOURCE	HTC; CAP trapper.
ORGANISM	Mus musculus (strain:C57BL/6J) 10 day old male pancreas cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library clone:1810060D19.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	Carninci,P. and Hayashizaki,Y.
JOURNAL	High-efficiency full-length cDNA cloning
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)
PUBMED	99279253
REFERENCE	10349636
AUTHORS	2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (1), 1617-1630 (2000)
MEDLINE	20495374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitunai,T., Tashiro,H., Itoh,M., Suni,I., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsuno,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujikawa,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matshiki,M., Yoneeda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	Kawai,J., Shingawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Aizawa,K., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Salto,T., Okazaki,Y., Gotohori,T., Bono,H., Kasukawa,T., Salto,R., Kadot,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochwa,H., Kusih,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schiml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Fletcher,P., de Bonaldo,M.F., Brownstein,M.J., Bulic,C., Carninci,P., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hotmann,M., Hume,D.A., Kamijima,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mommaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodrigues,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilmink,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kontsurk,S. and Hayashizaki,Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851
REFERENCE	5 (bases 1 to 772)
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Aizawa,K., Baldarelli,R., Bono,H., Brownstein,M., Bulic,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Horii,F., Hume,D., Imotohi,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koysa,S.,

TITLE  
JOURNAL

COMMENT  
Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayaishizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAAGACGCCAAGCCAGACCCTCTTTTCTTTTCTTTTN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 20.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGCGGGCCGCATATTTCGTGAGTAATTAAATCAATCCCCCCC 3'], cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.  
Location/Qualifiers  
1..772  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="FANTOM.DB:1810060D19"  
/db\_xref="MGD:MGI:1897650"  
/db\_xref="taxon:10090"  
/clone="1810060D19"  
/sex="male"  
/tissue\_type="pancreas"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="10 day old"  
1..772  
/gene="Tmem4"  
104..652  
/gene="Tmem4"  
/note="data source:MGD, source key:MGI:1928477,  
evidence:ISS  
putative  
transmembrane protein 4"  
/codon\_start=1  
/protein\_id="BAB25346.1"  
/db\_xref="GI:12841774"  
/db\_xref="MGD:MGI:1928477"  
/translation="MKGMWLALLGLGTANARRSODLHGCGRALVDLEMEIARVDRKTIOMGSFRINPDGOSGVYPARSEAHITELLEVCBMRKYGQIDPSTHRKNVRYVRNGESESLDIGIRIDSISGLTFKACESIVVEYDELLIEFSREADNKKDKTSKRTDICDLHAHSMDL"

BASE COUNT  
ORIGIN  
210 a 180 c 225 g 157 t

Alignment Scores:  
Score: No.:  
Percent Similarity: 98.35% Matches: 772  
Best Local Similarity: 97.25% Conservative: 177  
Query Match: 97.71% Mismatches: 2  
DB: 11 Indels: 3  
Gaps: 0

US-10-082-502-17 (1-182) x AK007914 (1-772)

1 Metlysglytrpgclgtrpleualaleuendcuglvalaleuencglthrtalatrpaala 20



```
Db 104 ATGAAGGCTGGGGTTGGCTAGCCCTACTTTTGGGGTCTGCTGGAACTGCTGGCT 163
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrp 40
Db 164 CGAAGGAGCCAGATCTACACTGTGAGGCTGTGAGGGCTGTGTGATGATTAAGTGG 223
QY 41 GlnIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60
Db 224 GAAATTGCCCGGTGGAGCCCAAGAACCATTCACATGGATGCCCTCCGAATCAATCCA 283
QY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
Db 284 GATGGAGCCAGCTACGTTGTGAGGACTTATGCCGCTCAGAGGCCACCTCACAGAG 343
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100
Db 344 TTGCTTGAGGAGGTGTGTACCGAATGAAGAGTACGGGGAGACAGATTGACCCCTTCTAC 403
QY 101 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyGluSerSerGluLeuAspLeu 120
Db 404 CACCCGCAAGAACTACGTACGCTCGTAGCCCGGAATGGGAATCCAGTGAACCTAGACTTAA 463
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 464 CAGGGCATCCGAATTGACTCAGATATCAGCGGCACCCCAAGTTGCTGTGAGAGCAATT 523
QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160
Db 524 GTGGAGAAATACGAGGATGAGCTTATCGAATTCCTCTCCAGAGAGGCTGACACGTTAA 583
QY 161 AspLysLeuCySerLysArgThrAspLeuCyAspHisAlaLeuHisIleSerHisAsp 180
Db 584 GACAAACTTTGCAGTAACGGAGACAGATCTATGTGACATGCCCTGACACAGATTCACGAT 643
QY 181 GluLeu 182
Db 644 GAGCTA 649
```

Search completed: March 31, 2003, 04:51:05  
Job time : 1173.85 secs



GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:24:04 ; Search time 1725.3 Seconds  
(without alignments)  
3070.029 Million cell updates/sec

Title: US-10-082-502-17  
Perfect score: 960  
Sequence: 1 MKGGMWLLALIGALGTAWA.....LCKSRKTDLDHALHSHDEL 182

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV-xlp  
-MODEL=frame+ \_p2n.model -DB=xlp  
-O=/cgn2.1/USPTO\_spool/US10082502/runat.24032003.135059.6547/app.query.fasta.1.1308  
-DB=GenEmbl -GFWT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biomum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=tbl -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10082502@cgn.1.1.4758\_@runat.24032003.135059.6547 -NCPU=6 -ICPU=3  
-NO\_XLPPX -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:.\*  
1: gb\_da:.\*  
2: gb\_hg:.\*  
3: gb\_in:.\*  
4: gb\_om:.\*  
5: gb\_ov:.\*  
6: gb\_pat:.\*  
7: gb\_ph:.\*  
8: gb\_pl:.\*  
9: gb\_pr:.\*  
10: gb\_ro:.\*  
11: gb\_sts:.\*  
12: gb\_sy:.\*  
13: gb\_un:.\*  
14: gb\_vl:.\*  
15: em\_ba:.\*  
16: em\_fun:.\*  
17: em\_hum:.\*  
18: em\_in:.\*  
19: em\_mu:.\*  
20: em\_om:.\*  
21: em\_or:.\*  
22: em\_ov:.\*  
23: em\_pat:.\*  
24: em\_ph:.\*  
25: em\_pl:.\*  
26: em\_ro:.\*  
27: em\_sts:.\*  
28: em\_un:.\*

29: em\_vl:.\*  
30: em\_hg\_hum:.\*  
31: em\_hg\_inv:.\*  
32: em\_hg\_other:.\*  
33: em\_hg\_mus:.\*  
34: em\_hg\_pln:.\*  
35: em\_hg\_rtd:.\*  
36: em\_hg\_mam:.\*  
37: em\_hg\_vrt:.\*  
38: em\_sy:.\*  
39: em\_hggo\_hum:.\*  
40: em\_hggo\_mus:.\*  
41: em\_hggo\_other:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		Query		DB	ID	Description
No.	Score	Match	Length			
1	960	100.0	814	6	AX440456	AX440456 Sequence
2	960	100.0	814	9	AB015631	AB015631 Homo sapi
3	960	100.0	1080	9	AY032624	AY032624 Homo sapi
4	960	100.0	1210	6	AX464040	AX464040 Sequence
5	938	97.7	798	10	BC008261	BC008261 Mus muscu
6	938	97.7	1017	10	AF186115	AF186115 Mus muscu
7	938	97.7	158198	2	AC012013	AC012013 Homo sapi
8	488.5	50.9	155023	2	AC025574	AC025574 Homo sapi
9	481	50.1	182350	2	AC073896	AC073896 Homo sapi
10	481	50.1	184762	2	AC023500	AC023500 Homo sapi
11	446.5	46.5	207424	10	AC090489	AC090489 Genomic S
12	436	45.4	101666	2	AC109891	AC109891 Rattus no
13	369	38.4	649	9	AF186113	AF186113 Homo sapi
14	328.5	34.2	69753	6	AC001027	AC001027 Homo sapi
15	266	27.7	564	6	AC103156	AC103156 Rattus no
16	246	25.6	2378	9	AX440375	AX440375 Sequence
17	246	22.1	996	3	AK094445	AK094445 Homo sapi
18	212	22.1	996	3	AY094698	AY094698 Homo sapi
19	190.5	19.8	40824	3	U13070	U13070 Caenorhabd
20	155.5	16.2	622	5	AB055671	AB055671 Arabidops
21	154	15.3	175963	3	AC009379	AC009379 Drosophill
22	147	15.3	279530	3	AE003518	AE003518 Drosophill
23	147	15.3	279530	3	AY060987	AY060987 Drosophill
24	144	15.0	1286	3	AC012854	AC012854 Drosophill
25	141	14.7	55359	2	AC007414	AC007414 Drosophill
26	141	14.7	172904	3	AE003831	AE003831 Drosophill
27	141	14.7	275390	3	AC008060	AC008060 Homo sapi
28	134.5	14.0	161566	9	AR142811	AR142811 Sequence
29	131.5	13.7	744	6	AX136616	AX136616 Sequence
30	131.5	13.7	746	6	BD006701	BD006701 Novel pol
31	131.5	13.7	747	6	BC019903	BC019903 Homo sapi
32	131.5	13.7	1413	9	AX136423	AX136423 Sequence
33	131.5	13.7	1419	6	AX376412	AX376412 Sequence
34	131.5	13.7	1470	6	AX454716	AX454716 Sequence
35	131.5	13.7	1470	6	AX491194	AX491194 Sequence
36	131.5	13.7	1512	6	AR142810	AR142810 Sequence
37	131.5	13.7	1512	6	BC032339	BC032339 Homo sapi
38	131.5	13.7	1547	9	AC015219	AC015219 Drosophill
39	129.5	12.5	76187	2	AC009403	AC009403 Homo sapi
40	122	12.7	168418	2	AC114126	AC114126 Rattus no
41	115	12.0	181896	2	AC120813	AC120813 Rattus no
42	115	12.0	201649	2	AC120813	AC120813 Rattus no
43	114.5	11.9	201649	2	AX380553	AX380553 Sequence
44	113	11.8	1260	6	AC129184	AC129184 Mus muscu
45	111	11.6	210631	2		

ALIGNMENTS

RESULT 1

AX440456 AX440456 814 bp DNA linear PAT 28-JUN-2002  
 LOCUS Sequence 309 from Patent WO0190154.  
 ACCESSION AX440456  
 VERSION AX440456.1 GI:21665266  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 AUTHORS 1 Xu, J., Mitchem, J. L., Harlocker, S. L., Dillon, D. C., Secrist, H.,  
 Lodes, M. J., Algate, P. A., Pling, S. P., Mannion, T., Benson, D. R. and  
 Carter, D.  
 TITLE Compositions and methods for the therapy and diagnosis of ovarian  
 cancer  
 JOURNAL Patent: WO 0190154-A 309 29-NOV-2001;  
 CORIXA CORPORATION (US)  
 FEATURES  
 SOURCE 1. 814  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 210 a 194 c 241 g 169 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.15e-87 Length: 814  
 Score: 960.00 Matches: 182  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-10-082-502-17 (1-182) x AX440456 (1-814)

Oy 1 Mettysglttprglttprleuallaleuueuglualaleuueuglthralatrrala 20  
 |||||  
 Db 145 ATGAAAGCTGGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 204  
 Oy 21 Argatserglnasrpleuallaleuueuglualaleuueuglthralatrrala 40  
 |||||  
 Db 205 CGAGAGACCGAGATCTCCAGTGGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAG 264  
 Oy 41 Glutleallaglnvalasrprolysthrllleglmetylserphearglleaspro 60  
 |||||  
 Db 265 GAAATGGCCAGGTGGAGCCCAAGAGACCATTCAGATGGATCTTCCGATCAATCA 324  
 Oy 61 AspglyserglnserValalgluValaProtyrAlaArgserglualaleuueuglth 80  
 |||||  
 Db 325 GATGGACACGATCGATGGTGGAGGTGGCTTATGCCCTCAGAGGCCACCTCAGAGG 384  
 Oy 81 Leuleuglulaleucysarpargmetlysglultrglulglulnllaasrproserthr 100  
 |||||  
 Db 385 CTGCTGGAGGAGATGTGACCGGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 444  
 Oy 101 HlsargluysaenttyValalgluValaLysarglansgllylserSergluleuasp 120  
 |||||  
 Db 445 CATCCCAAGAACTAGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTAC 504  
 Oy 121 GlnsllyleatrglleasrserasrllleserlythleuysrphelaCysglulserlle 140  
 |||||  
 Db 505 CAAAGCATCCGATGATCGATGATGATGATGATGATGATGATGATGATGATGATGATG 564  
 Oy 141 Valglulutrglulaspgluleuileglurpheraserarglualaleuueuglth 160  
 |||||  
 Db 565 GTGGAGAAATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 624  
 Oy 161 Asprlyleucysrserlysthrllleglmetylserphearglleaspro 180  
 |||||  
 Db 625 GACAAACTTTCAGTAAACGAGATCTTGTGACATGCGCTGCACATATCGCATGAT 684  
 Oy 181 Gluleu 182  
 |||||

Db 685 GAGCTA 690  
 RESULT 2  
 LOCUS AB015631  
 DEFINITION Homo sapiens mRNA for type II membrane protein, complete cds,  
 clone:HP10390.  
 ACCESSION AB015631  
 VERSION AB015631.1 GI:4586839  
 KEYWORDS type II membrane protein.  
 SOURCE Homo sapiens gastric adenocarcinoma CDNA to mRNA,  
 clone:lib:pkal-meta-1 clone:HP10390.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 AUTHORS 1 Yokoyama-Kobayashi, M., Yamaguchi, T., Sekine, S. and Kato, S.  
 TITLE Selection of cDNAs encoding putative type II membrane proteins on  
 the cell surface from a human full-length cDNA bank  
 JOURNAL Gene 228 (1-2), 161-167 (1999)  
 MEDLINE 99173880  
 REFERENCE 2 (bases 1 to 814)  
 AUTHORS Kato, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-JUN-1998) Satshi Kato, Research Institute of National  
 Rehabilitation Center for the Disabled, Department of  
 Rehabilitation Engineering, 4-1 Namiki, Tokorozawa, Saitama  
 359-8555, Japan (E-mail:setsu@rehab.go.jp).  
 Tel:042-995-3100(ex.2568), Fax:042-995-3133)  
 FEATURES  
 SOURCE 1. 814  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="HP10390"  
 /tissue="gastric adenocarcinoma"  
 /clone\_lib="pkal-meta-1"  
 145. 693  
 /product="type II membrane protein"  
 /protein\_id="PAA76498.1"  
 /db\_xref="gi:4586840"  
 /translation="MKGGKVALLLGALLGTRMKRRSDODLHCGACRLVDELEETIQ  
 VPKKTIOMGSRINPDSQSVVEVPAKRSVAHLELEECIDMKKKEGQIDSTIR  
 KNYRVVGRNSESSELDQGRIDSDISDTLKFACESIVEREDELIEFSREADNVK  
 DKLSKRTDLCDHALHSHDEL"

US-10-082-502-17 (1-182) x AB015631 (1-814)

Oy 1 Mettysglttprglttprleuallaleuueuglualaleuueuglthralatrrala 20  
 |||||  
 Db 145 ATGAAAGCTGGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 204  
 Oy 21 Argatserglnasrpleuallaleuueuglualaleuueuglthralatrrala 40  
 |||||  
 Db 205 CGAGAGACCGAGATCTCCAGTGGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAG 264  
 Oy 41 Glutleallaglnvalasrprolysthrllleglmetylserphearglleaspro 60  
 |||||  
 Db 265 GAAATGGCCAGGTGGAGCCCAAGAGACCATTCAGATGGATCTTCCGATCAATCA 324  
 Oy 61 AspglyserglnserValalgluValaProtyrAlaArgserglualaleuueuglth 80  
 |||||  
 Db 325 GATGGACACGATCGATGGTGGAGGTGGCTTATGCCCTCAGAGGCCACCTCAGAGG 384

QY	81	LeuLeuGIUGUILLiLeuSAspArpIcmetLysGIuYUTrGIUGUInLIAspProSerThr	100
Db	385	CTGCTGGAGGAGATATGTGTGACCGGATGAAAGAGTATGGGGAACAGATGATCTTCCACC	444
QY	101	HisArgLysAspTyrValArgValArgValArgAsnGIuSerSerGIuLeuAspLeu	120
Db	445	CATGCGCAAGAACTACGTACGTGTAGTGGCGCGGATGGAAATCCAGTGAATCGGACCTTA	504
QY	121	GlnGIUllLeaArgIIeAspSerAspIIeSerGIUTrLeuLysPheAlaCysGIuSerIle	140
Db	505	CAAGCAATCCCAATCCAGTCAAGATATGTACGGCACCCCTCAAGATTGGCTGTGACAGCAATT	564
QY	141	ValGIUGUTrGIUAspGIuLeuIIeGIUTrPhePheSerArgGluAlaAspAsnValLys	160
Db	555	GTGGAGGAATACGAGAGATGAATCTATTAAATCTTTTCCCGAGAGGCTGACAAATGTAA	624
QY	161	AspLysLeuCysSerLysArgTrThrAspLeuCysAspAsnIAlaLeuAlaIIeSerHisAsp	180
Db	625	GACAACTTTTGCAGTAAGCGAAACAGATCTTTGTGACGATGGCCCTGCACATATCGCATGAT	684
QY	181	GIuLeu 182	
Db	685	GAGCTA 690	
RESULT 3			
AY032624			
LOCUS	AY032624	1080 bp	mRNA linear PRI 18-APR-2002
DEFINITION	Homo sapiens saposin-like protein mRNA, complete cds.		
ACCESSION	AY032624		
VERSION	AY032624.1	GI:20196198	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1080)		
TITLE	Bornhauser,B.C., Olsson,P.-A. and Lindholm,D.		
JOURNAL	NSAP is a novel saposin-like protein that interacts with MIR and stimulates neurite outgrowth		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 1080)		
TITLE	Olsson,P.-A. and Lindholm,D.		
JOURNAL	Direct Submission		
FEATURES	Submitted (13-APR-2001) Neuroscience, Uppsala University, Husergatan 3, Uppsala 75123, Sweden		
source	Location/Qualifiers		
	1..1080		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="12"		
	/map="12q15"		
	419..367		
	/note="NSAP; encodes a type 2 membrane signal and a putative ER retention motif"		
	/codon_start=1		
	/product="saposin-like protein"		
	/protein_id="AAK38148.1"		
	/db_xref="GI:20196199"		
	/translation="MKMGWLLALIGLCTGWAARSQDLHCACRALVDELLEWELAQVDPKLTOMGSFRINPDGSSVVEVPARSAHNTLEELICDPMKEGEOIDSTFRHKNYRVVGRNGSESELDLGIIRIDSDISGLTKFACESIVEVEDELLIEFREADNWVKDKLCKSRDLCNDHLNHTSHDEL"		
BASE COUNT	238 a 269 c 343 g 230 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1,64e-87	Length:	1080
Score:	960.00	Matches:	182
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

<hr/>					
US-10-082-502-17	(1-182)	x	AY032624	(1-1080)	
OY	1	MellysglTyrpgjUTrPLEuAlALeuleuDeuglAlAleuleuGlUThrAlATrPA	20		
Dd	419	ATGAAAGCGCTGGGTGGCTGCCCTGGCTTGCGGGGGCCCTGCCTGGAAACCGCGCTGGCT	478		
OY	21	ArgArGSerGlnArlPleuHnIscGsgLUlaAcysArGalAlaleuValAspGluDeuGLUTrr	40		
Dd	479	CGGAGGACCSCGAGATCTCCACTGTGGACCAACAGGGCTGTGGATGAACCTGAATGG	538		
OY	41	GluIUleIAgInVAlAspProLyLeuThrILegImetGlySerPharGLIleasnPro	60		
Dd	539	GAATTGCCCGAGTGGACCCCCAAGAACCATTCCAGATGGAGTCTTCGGATCCAATCA	598		
OY	61	AspGLSerGlnSerValAlGIuValProTUrAlArGserGluAlAnIslaueThrgLu	80		
Dd	599	GATGGCACCCAGTCAGTGTGTGAGGTGCTTAATGCCCTCAGAACGCCCATCTACAAG	658		
OY	81	LeuIeuGLUGluIIIECyAsArArgmetLysGluTyrgLUGluGlnIIleasProSerThr	100		
Dd	659	CTGCGAGGAGAAATGTGACCCGGATGAAGAGTAGTGGGAACAAATGTGATCTCCACCC	718		
OY	101	HIsArGLySAsnTYrValArGLValAlGIuTyrgAsnGLyGluSerSerGluDeuAspLeu	120		
Dd	719	CATGCCAAGAACTAGTACGTCTGTACTGGCCCGGAAATGAGAAATCCAGTGAACCTGACCTTA	778		
OY	121	GlnGLyIleArGLleAsPSerAspRIleSerGLyThreULeULysrPhalaCySGluSerIle	140		
Dd	779	CAAGCATCCCAAATCGACTCAGATATGTAGCGGACCCCTCAAGTTGTCGTGTGAGAGCATTT	838		
OY	141	VAlGluGluTyrgUAlAspRGLIuIleuIIIEIuherPheSerArGLuAlaAspAsnValLys	160		
Dd	839	GTGGAGAAATCGAGAGATGAACTATTGAATCTTTTCCGAGAGCGTGAATAATGTTAAA	898		
OY	161	AspLysLeuCysSerLYArGLThrAspLeuCysAspRIAlaLeuHnIslaueThrgIleSerHIsasp	180		
Dd	899	GACAAACTTTTCCAGTAACGAAACGAAAGATTTTGTGTGACATCGCCCTGCACATATCCCATAT	958		
OY	181	GLuLeu 182			
Dd	959	GAGCTA 964			
<hr/>					
RESULT 4					
AX464040					
LOCUS	AX464040	1210 bp	DNA	linear	PAT 16-JUL-2002
DEFINITION	Sequence 173 from Patent WO0140466.				
ACCESSION	AX464040				
VERSION	AX464040.1	GI:21899037			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
TITLE	Baker,K.P., Bershtal,M., DeForge,L., Desnoyers,L., Filveroff,E.,				
JOURNAL	Geo,W.Q., Gerritsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,				
FEATURES	Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,				
SOURCE	Wood,W.L. and Zhang,Z.				
BASE COUNT	Patent: WO 0140466-A 173 07-JUN-2001;				
ORIGIN	Genentech Inc. (US)				
	Location/Qualifiers				
	1..1210				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	261 a 316 c 379 g 254 t				
<hr/>					
Alignment Scores:	1.88e-87				1210
Fred. No.:	960.00				Matches:
Score:					182

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-082-502-17 (1-182) x AX464040 (1-1210),

OY	1	Melysglytprglttpttleualaleuueuglvalaleulemgltyhrllatrpala	20
Db	543	ATGAAGGCTGGCGTGGCTGGCCCTGCTTTCTGGGGCCCCCTTCGTGGAAACCCCTGGCGCT	602
OY	21	ArgArSerGIaaRLeuHnIScysGLValAcysArGalaleuValAspLubdeugltpr	40
Db	603	CcGAgAGGCCAGSGAtTCTCActGTGTGACACTGCAGAGCGCTGTGTGATGAATAGAATGG	662
OY	41	GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheargIleasnpro	60
Db	663	GAATTGGCCAGGTGGACCACCAAGAACCAATCAGATGGGATCTTTCGGGTACAATCCA	722
OY	61	AspGlySerGInserValValGlValPrOTryAlaATrSerGluAlaHISLeuThrglu	80
Db	723	GATGcGACGCCAGTCAGTGCTGTGGAGGTGCCCTTAATGCCCTCGAGAGCCCACTCACAGAG	782
OY	81	LeuLeuGluGluIleCysAspArgMetySglutryGlyGluGlnIleAspProSerthr	100
Db	783	CTGTGTGGAGGAATGATGACCGGCTGAAGAGATGTGGGGAAACAATTCATCTTCCACCC	842
OY	101	HISArgLysAaNTryValArgValValGlTyrrAsnGlyGlyUSeSerGluLeuAspLeu	120
Db	843	CATGCGAAGAACTACGTACGTGTAGTGGGCCGGAATGGAGAAATCCAGTAGTACGTGACCTA	902
OY	121	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerile	140
Db	903	CAAGGCAATCCCAAATCGACTCAGATATGTAGCCGACCCCAAGTTTGCCTGTAGAGCATT	962
OY	141	ValGluGluTryGluAspGluIleGluPhePheSerArgGluAlaAspAsnVallys	160
Db	963	GTGAGAGAAATCGAGAGATGAACATCTTAATTTCTTTCCCAGAGAGCTGACATGTTTAA	1022
OY	161	AspLysLeuCySerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisasp	180
Db	1023	GACAAACTTTTCAATAAGCGAAGACAGATCTTTGTGCACATGCCCTGCACATATCGCATGAT	1082
OY	181	GluLeu 182	
Db	1083		
		GAGCTA 1088	
RESULT 5			
BC008261			
LOCUS		798 bp mRNA linear ROD 07-AUG-2002	
DEFINITION		Mus musculus, transmembrane protein 4, clone MGC:6853	
ACCESSION		IMAGE:2650612, mRNa, complete cds.	
VERSION		BC008261	
KEYWORDS		BC008261.1 GI:14198400	
SOURCE		MGC.	
ORGANISM		house mouse.	
		Mus musculus	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE		1 (bases 1 to 798)	
AUTHORS		Straussberg,R.	
TITLE		Direct Submission	
JOURNAL		Submitted (22-MAY-2001) National Institutes of Health, Mammalian	
		Gene Collection (MGC), Cancer Genomics Office, National Cancer	
		Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,	
		USA	
REMARK		NIH-MGC Project URL: http://mgc.ncl.nih.gov	
COMMENT		Contact: MGC help desk	
		Email: cgabbs-r@mail.nih.gov	
		Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.	
		cDNA Library Preparation: Life Technologies, Inc.	
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)	
		DNA Sequencing by: Baylor College of Medicine Human Genome	

Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: amg@bcm.tmc.edu  
Gunnarntne, P.H., Garcia, A.M., Lu, X., HuiJyk, S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.

Clone distribution: MCC clone distribution information can be found  
through the I.M.A.G.E. Consortium/INL at: <http://image.lnl.gov>  
Series: IRAC Plate: 5 Row: b Column: 23  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 9903606.

location/Qualifiers  
1..798

**CDS**

BASE COUNT	227 a	182 c	228 g	161 t
ORIGIN				

Alignment Scores:	
Pred. No.:	1.84e-85
Score:	938.00
Percent Similarity:	98.35%
Best Local Similarity:	97.25%
Query Match:	97.71%
DB:	10
Length:	796
Matches:	177
Conservative:	2
Mismatches:	3
Indels:	0
Gaps:	0

US-10-082-502-17 (1-182) x BC008261 (1-798)

Qy	1	Metysylslyrpsqlyltyrpleuvalaleuileuencylvalaleuileuglythrallatrpala	20
Db	115	ATGAAAGCGTGGGGTGTGGCTAGACCCCTACTTTGGGGGCTCCTCGGGAACTGGCTGGGCT	174
Qy	21	ArgArGserGlnaspIeuHlsCYsGlyAlaCYsarGalaLeuValaAspGluLeuGluTrp	40
Db	175	CGAAGGAGCCAAAGTCTCAACACTGTGGAGCTTGGCGGGCTCTGTGGTGTGATTAGAGTGG	233
Qy	41	GluIleuLeuGlnValaAspProLysLyThrIleGlnMetGlySerPheArgIleAsnPro	60
Db	235	GAATTTCGCCCGCGTGGACCCCAAGAACCAATTACATGGGATCTTCCGAATCAATCA	294
Qy	61	AspGlySerGlnSerValValGluValProTyralaArgserGluAlaHlsLeuThrGlu	80
Db	295	GATGGCAGCCAGTCACTGTTGTGGAGGTACCTTATCCCGCTCAGAGGCCCACTCACAGAG	354
Qy	81	LeuIeuGlnGluIleCYsaAspArgMetLysGluTrpGlyGlnIleAspProSerThr	100
Db	355	TTGGTTAGAGGAGGTGTGACCGAATACAGAGTACGGGGAACAGATTGACCTTCAACC	414
Qy	101	HlsArgLysAsnTyrlaArgvalValGlyArgsnGlyLysSerGluLeuAspLeu	120
Db	415	CACCGCAAGAACACTACGTAACGCGTCGTGAGCGGAATGCAAAATTCAGTAACACTGACTTA	474
Qy	121	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCYsgluSerIle	140

Db 475 CAGGCATCCGAATTCAGATATACAGCGCACCTCAGTTCCGTGTGAGACATT 534  
 QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnVallys 160  
 Db 535 GTGGAGAAATACAGAGATAGCTATCGAATTTCTCCAGNAGGCTGCACAGTTAA 594  
 QY 161 AspGluLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisasp 180  
 Db 595 GACAAACTTTCAGTAAGCGGACAGATCTATGTGACCATCCTCGACAGATCTCAGAT 654  
 QY 181 GluLeu 182  
 Db 655 GAGCTA 660  
 RESULT 6  
 AF186115 1017 bp mRNA linear ROD 13-JAN-2000  
 LOCUS Mus musculus putative secreted protein Zs1g9 (Zs1g9) mRNA, complete  
 DEFINITION cds.  
 ACCESSION AF186115 GI:6014635  
 VERSION AF186115.1  
 KEYWORDS Mus musculus.  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 1017)  
 AUTHORS Sheppard, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J. and  
 O'Hara, P.  
 TITLE Mus musculus putative secreted protein  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1017)  
 AUTHORS Sheppard, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J. and  
 O'Hara, P.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-SEP-1999) Biomolecular Informatics, ZymoGenetics,  
 Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA  
 FEATURES  
 source  
 1..1017  
 /organism="Mus musculus"  
 /db\_xref="dbEST:AA049839"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:476061"  
 1..1017  
 /gene="Zs1g9"  
 /gene=".891"  
 /product="putative secreted protein Zs1g9"  
 /protein\_id="AA01433.1"  
 /db\_xref="GI:6014636"  
 /translation="MKGMGLALLGLLTGTAMARRSODLHGACRALVDELEMTAR  
 VDPKTIOMGSRINPDGSOVVEYVAEHLLELLEVDCKMEYEEODIPSHR  
 KNYRVSRNGSSSELIDGICIRIDISITLKFACISYEVEDELEIEFFSRADNVK  
 DLKSKRTLDLCHALHRSHDEL"  
 CDS  
 gene  
 343..891  
 /gene="Zs1g9"  
 /product="putative secreted protein Zs1g9"  
 /protein\_id="AA01433.1"  
 /db\_xref="GI:6014636"  
 /translation="MKGMGLALLGLLTGTAMARRSODLHGACRALVDELEMTAR  
 VDPKTIOMGSRINPDGSOVVEYVAEHLLELLEVDCKMEYEEODIPSHR  
 KNYRVSRNGSSSELIDGICIRIDISITLKFACISYEVEDELEIEFFSRADNVK  
 DLKSKRTLDLCHALHRSHDEL"  
 BASE COUNT 248 a 250 c 315 g 204 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,49e-85 Length: 1017  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3  
 Query Match: 97.71% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-082-502-17 (1-182) x AF186115 (1-1017)  
 QY 1 MetLysGlyTyrPglYrPleuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20  
 Db 343 ATGAAGGCGGGGTTGGCTTCTTTGGGGGCTCTGCGGGAATCGCTGGGCT 402  
 QY 21 ATGATGSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyrP 40

Db 403 CGAAGAGCCCAAGATCTACAGCTGTGAGCTTCGAGGCGCTGTGTCGATCAATAGAGTGG 462  
 QY 41 GluIleAlaGluValAspProLysLysThrIleGluMetGlySerPheArgIleAsnPro 60  
 Db 463 GAATTTGCCCGCGTGGAGCCCAAGAAAGACCATTCAATGGATGCTTCGAAATCAATCCA 522  
 QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
 Db 523 GATGGCAGCAGTCACTTTGTGAGGTACCTTATGCCCGCTCAGNAGGCCACCTCAGAG 582  
 QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
 Db 583 TTTCCTTACAGAGAGGTGTGTGACCGAATGAAGAGTACGGGCAACAGATTACCTTAC 642  
 QY 101 HisArgLysAspTyrValArgValValGlyArgArgGlyGlnSerSerGluLeuAspLeu 120  
 Db 643 CACCGCAAGAACTACATACGTCGTGAGCCGGAATGGAATCCAGTGAATCACTTA 702  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlnSerIle 140  
 Db 703 CAGGCATCCGATTCAGTACTAGATATCAGCGGACCTCAAGTTTCGTGTGAGACATT 762  
 QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnVallys 160  
 Db 763 GTGGAGAAATACAGAGATAGCTATTCGAATTTCTCCAGAGAGCTGCACAGTTAA 822  
 QY 161 AspGluLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisasp 180  
 Db 823 GACAAACTTTCAGTAAGCGGACAGATCTATGTGACCATCCTCGACAGATCTCAGAT 882  
 QY 181 GluLeu 182  
 Db 883 GAGCTA 888  
 RESULT 7  
 AC012013 158198 bp DNA linear HTG 17-JUL-2001  
 LOCUS Homo sapiens chromosome 12 clone RP11-764L14, WORKING DRAFT  
 DEFINITION SEQUENCE, 5 unordered pieces.  
 ACCESSION AC012013 GI:14717292  
 VERSION AC012013.18  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 158198)  
 AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,  
 Alstbrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbarta, J.,  
 Benton, J., Blinze, K., Blankenburg, K., Bonnin, D., Bouck, J.,  
 Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,  
 Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,  
 Carter, N., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,  
 Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C.,  
 Coyle, M.D., Dalhorne, S.R., David, R., Davila, M.L., Davis, C.,  
 Davy-Carroll, L., Dederich, D.A., Delaney, R.R., Delgado, H.,  
 Dem, A.L., Ding, Y., Din, H.H., Douthwaite, K.J., Draper, H.,  
 Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,  
 Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,  
 Foster, P., Frantz, P., Gablitz, A., Gao, J., Garcia, A., Garner, T.,  
 Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S.,  
 Hamilton, K., Harris, C., Harris, K., Hart, M., Havlik, P., Hayes, A.,  
 Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C.,  
 Hollins, B., Homsl, F., Howard, S., Huber, J., Hulys, S., Hume, J.,  
 Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,  
 Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,  
 Kovah, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,  
 Lewis, L., Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W.,  
 Louised, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
 Ma, J., Maheshwari, M., Mapue, P., Martin, R., Martindale, A.,  
 Martinez, E., Massey, E., Mawliny, E., McLeod, M.P., Meador, M.,  
 Mel, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,

Alignment Scores:	1,31e-82	Length:	158198
BASE COUNT	46057 a	30922 c	31747 g 49030 t 442 others
ORIGIN			
FEATURES			
SOURCE	<p>1. 158198</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="taxon:9606"</p> <p>/chromosome="12"</p> <p>/clone="Rp11-764L14"</p>		
REFERENCE	<p>108229: contig of 108229 bp. in length</p> <p>108230 150246: gap of unknown length</p> <p>108330 150246: contig of 41917 bp in length</p> <p>150247 150346: gap of unknown length</p> <p>150347 153410: contig of 3064 bp in length</p> <p>153411 153510: gap of unknown length</p> <p>153511 155739: contig of 2229 bp in length</p> <p>155740 155839: gap of unknown length</p> <p>155840 158198: contig of 2359 bp in length.</p>		
COMMENT	<p>***** Genome Center *****</p> <p>Center: Baylor College of Medicine</p> <p>Center code: BCM</p> <p>Web site: <a href="http://www.hgsc.bcm.tmc.edu/">http://www.hgsc.bcm.tmc.edu/</a></p> <p>Contact: <a href="mailto:hgsc-help@bcm.tmc.edu">hgsc-help@bcm.tmc.edu</a></p> <p>----- Project Information -----</p> <p>Center project name: RP11-764L14</p> <p>Center clone name: RP11-764L14</p> <p>----- Summary Statistics -----</p> <p>Sequencing vector: M13; L08821</p> <p>Chemistry: Dye-terminator Big Dye; 100% of reads</p> <p>Assembly program: Phrap; version 0.990329</p> <p>Consensus quality: 161278 bases at least Q40</p> <p>Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,</p> <p>Consensus quality: 174843 bases at least Q20</p> <p>Estimated insert size: 168803; sum-of-contigs estimation</p> <p>Estimated insert size: 163288; agarose-1p estimation</p> <p>Quality coverage: 6x in Q20 bases; agarose-1p estimation</p> <p>Quality coverage: 5.8x in Q20 bases; sum-of-contigs estimation</p> <p>*****</p> <p>* NOTE: Estimated insert size may differ from sequence length</p> <p>* (see <a href="http://www.hgsc.bcm.tmc.edu/docs/genbank_data.html">http://www.hgsc.bcm.tmc.edu/docs/genbank_data.html</a>)</p> <p>* NOTE: This is a 'working draft' sequence. It currently</p> <p>* consists of 5 contigs. The true order of the pieces</p> <p>* is not known and their order in this sequence record is</p> <p>* arbitrary. Gaps between the contigs are represented as</p> <p>* runs of N, but the exact sizes of the gaps are unknown.</p> <p>* This record will be updated with the finished sequence</p> <p>* as soon as it is available and the accession number will</p> <p>* be preserved.</p>		
TITLE	<p>2 (bases 1 to 158198)</p> <p>Unpublished</p> <p>Direct Submission</p> <p>Worley,K.C.</p> <p>Submitted (19-Oct-1999) Human Genome Sequencing Center, Department</p> <p>of Molecular and Human Genetics, Baylor College of Medicine, One</p> <p>Baylor Plaza, Houston, TX 77030, USA</p> <p>On Jul 13, 2001 this sequence version replaced gl:14547727.</p>		
JOURNAL	<p>*****</p> <p>Center: Baylor College of Medicine</p> <p>Center code: BCM</p> <p>Web site: <a href="http://www.hgsc.bcm.tmc.edu/">http://www.hgsc.bcm.tmc.edu/</a></p> <p>Contact: <a href="mailto:hgsc-help@bcm.tmc.edu">hgsc-help@bcm.tmc.edu</a></p> <p>----- Project Information -----</p> <p>Center project name: RP11-764L14</p> <p>Center clone name: RP11-764L14</p> <p>----- Summary Statistics -----</p> <p>Sequencing vector: M13; L08821</p> <p>Chemistry: Dye-terminator Big Dye; 100% of reads</p> <p>Assembly program: Phrap; version 0.990329</p> <p>Consensus quality: 161278 bases at least Q40</p> <p>Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,</p> <p>Consensus quality: 174843 bases at least Q20</p> <p>Estimated insert size: 168803; sum-of-contigs estimation</p> <p>Estimated insert size: 163288; agarose-1p estimation</p> <p>Quality coverage: 6x in Q20 bases; agarose-1p estimation</p> <p>Quality coverage: 5.8x in Q20 bases; sum-of-contigs estimation</p> <p>*****</p> <p>* NOTE: Estimated insert size may differ from sequence length</p> <p>* (see <a href="http://www.hgsc.bcm.tmc.edu/docs/genbank_data.html">http://www.hgsc.bcm.tmc.edu/docs/genbank_data.html</a>)</p> <p>* NOTE: This is a 'working draft' sequence. It currently</p> <p>* consists of 5 contigs. The true order of the pieces</p> <p>* is not known and their order in this sequence record is</p> <p>* arbitrary. Gaps between the contigs are represented as</p> <p>* runs of N, but the exact sizes of the gaps are unknown.</p> <p>* This record will be updated with the finished sequence</p> <p>* as soon as it is available and the accession number will</p> <p>* be preserved.</p>		
AUTHORS	<p>Morgan,M., Morris,S., Moser,M., Neali,D., Newton,J., Newton,N.,</p> <p>Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokweto,S.,</p> <p>Oguni,M., Okunubu,G., Oragunde,N., Oviedo,R., Pace,A., Payton,B.,</p> <p>Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,</p> <p>Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,</p> <p>Ruit,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshitari,N.,</p> <p>Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,</p> <p>Stone,H., Sutton,A., Svatek,A., Tabori,P., Tamerisa,A., Tamerisa,K.,</p> <p>Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,</p> <p>Thomas,S., Umanal,K., Vasquez,L., Vera,V., Willard,D., Vinson,R.,</p> <p>Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,</p> <p>Watlington,S., Williams,G., Williamson,A., Wleczky,K., Wooden,S.,</p> <p>Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,</p> <p>Wernstock,G. and Gibbs,R.</p>		

Percent Similarity:	938.00	Matches:	177
Score:	98.35%	Conservative:	3
Best Local Similarity:	97.25%	Mismatches:	2
Query Match:	97.71%	Indels:	0
Db:	2	Gaps:	0
US-10-082-502-17 (1-182) x AC012013 (1-158198)			
QY 1	MetylsGlyTTPbGlyTTPLeuAlaLeuLeuLeuGlyAlaLeuLeuGlyThAlaTTPAla 20		
Db 152111	ATGAAAGCGCTGGGGTGGCTTACGCTTACCTTTGGGGGCTCTGGGAAACGCTGGGCT 152170		
QY 21	ArgATSerGlnAspLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuGlnTTP 40		
Db 152171	CGAAGGACCGAAGATCTTACACTGTGGACCTTGCAGGGCTTGTGGATGATTAAGATGG 152230		
QY 41	GluIleAlaGlnValAspProLysPheLysThrIleGlnMetGlySerPheArgIleAsnPro 60		
Db 152231	GAATATGGCCGGGTGGACCCCAAGAAGACATTCAGATGGGATCCCTCCGAATCATGCCA 152290		
QY 61	AspGlySerGlnSerValAlaGluValProGlyAlaArgSerGluAlaHISLeuThrGlu 80		
Db 152291	GATGGCACCCAGTCAGTTGTGGAGGTACTTATGCCCCCTCAGAGGCCACCTCACAAG 152350		
QY 81	LeuLeuGluGluIleCysAspArgMetLysGluIlyrGlyGluGlnIleAspProSerThr 100		
Db 152351	TTGCTTGAGAGAGGTGTGGACCGAATGAAAGATGCGGGGAACAGATTGACCTTTCACC 152410		
QY 101	HISArgLysAsnTyrValArgValAlaGluValArgAsnGlyLysSerGluLeuAspLeu 120		
Db 152411	CACCCCAAGAACTACGTACGGGTGCTGTGACCCGGAATGAGAATCCAGTAACATGACTTA 152470		
QY 121	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140		
Db 152471	CAGGCAATCCCAATGTGACTCAGATATACGCGCACCCCTCAAGTTTGCCTGTGAACAGCAT 152530		
QY 141	ValGluGluIlyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160		
Db 152531	GTGAAAGATACGAGGATGACCTTTCGAATCTTCTCCAAAGAGGCTGACAAACGTTTAA 152590		
QY 161	AspLysLeuCysSerLysArgThrAspLeuCysAspHISAlaLeuHISLeSerHISAsp 180		
Db 152591	GACAAACCTTTCAGTAACCGGACAGATTTATGTGACCATGCGCCGCACACAGATCTCACGAT 152650		
QY 181	GluLeu 182		
Db 152651	GAGCTA 152656		
RESULT 8			
LOCUS	AC025574/c	155023 bp	DNA linear HTG 09-MAY-2002
DEFINITION	Homo sapiens chromosome 12 clone RP11-348M3, WORKING DRAFT		
ACCESSION	AC025574		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 155023)		
AUTHORS	Kuzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alshroob,S.L., Amaralunge,H.C., Are,J.R., Ayalew,M., Banks,T., Barakia,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Bukey,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,		



```

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_drfmt\_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
      1      3033: contig of 3033 bp in length

```

	*	3034	3133:	gap of unknown length	
	*	3134	5971:	contig of 2838 bp in length	
	*	5972	6071:	gap of unknown length	
	*	6072	10269:	contig of 4198 bp in length	
	*	10270	10369:	gap of unknown length	
	*	10370	13009:	contig of 2640 bp in length	
	*	13010	13109:	gap of unknown length	
	*	13110	17573:	contig of 4466 bp in length	
	*	13110	17573:	contig of 4466 bp in length	
	*	17674	24650:	contig of 6977 bp in length	
	*	24651	24751:	gap of unknown length	
	*	31779	31778:	contig of 7028 bp in length	
	*	31879	42263:	contig of 10865 bp in length	
	*	42264	42363:	gap of unknown length	
	*	42364	51105:	contig of 8741 bp in length	
	*	51105	51204:	gap of unknown length	
	*	51205	68306:	contig of 17102 bp in length	
	*	68307	68406:	gap of unknown length	
	*	68407	109481:	contig of 41075 bp in length	
	*	109482	109581:	gap of unknown length	
	*	109582	155023:	contig of 45442 bp in length.	
FEATURES					
source			1..155023		
		/organism="Homo sapiens"			
		/db_xref="taxon:9606"			
		/chromosome="12"			
		/clone="RP11-348M3"			
BASE COUNT	40008 a	36207 c	35987 g	41693 t	1128 others
ORIGIN					
Alignment Scores:					
Pred. No.:		2,31e-37	Length:	155023	
Score:		488.50	Matches:	114	
Percent Similarity:		39.72%	Conservative:	0	
Best Local Similarity:		39.72%	Mismatches:	0	
Query Match:		50.89%	Indels:	173	
DB:		2	Gaps:	2	
US-10-082-502-17 (1-182) x AC025574 (1-155023)					
OY	69	VAlProLyrrAArGserGUalAHIsLeuthrgLUleuGLugluILecyAsPArq			88
Db	56120	GTGGCTTATCCCGGCACAGGGGCCACCCTCACAGACTGTGGAGGATATGTGACCG			56061
OY	89	MeLysGIUrYrIGlUGInLIeAsPrOsErThrHSAdgLYsaNryrValArGval			108
Db	56060	ATGAAGAGATATGGGACACAGATTGATTCCTCCACCCCTCCCAAGAACCTACGCTGT			56001
OY	109	VAlGlYrGAnsIyGLUsErsErgLUleuAsPlEuIngLIyIArGIleAsPsrAap			128
Db	56000	GTGGGCCGGAAATGGAATCCAGTGAACTGACCTACACAAGCATCCGAATCGACTCAGAT			55941
OY	129	IleSerGIYThrLeuUlyssPhela-----			136
Db	55940	AATAGCGGCAACCTCAAGTTTCGGGTGAGCTATGGGAATCGTAGCTGTTCTTGGACAT			55881
OY	136	-----			136
Db	55880	TAAAGGGTGTGGAGAACCAATTGAAAGTTAGGCTGATATCTGTCTCTCTTTGG			55821
OY	136	-----			136
Db	55820	AAGTGAAGCAAGAGCGCTTTATTTGCCCTGTGCACCACTTCTCCCTTGGTTGGAG			55761
OY	137	CySGlUsErlEvaIgLUglULyGrIUnAsPlEuIleGIunPhEsErArGGuIlA			156
Db	55760	TGTGAAGAGCATTTGTGAGAGATACGAGAGATGACTATTAATTTCTTTCCGAGAGGCT			55701
OY	157	AsPaNsValYAsApLYsLeUCysSErLYArGthr-----			168
Db	55700	GACATGTTTAAAGACAACTTTGACATGACGAAGAGTAAAGCTGCCCCCACTTATATCTC			55641

```

Oy 168 ----- 168
Db 55640 CTGTTCTAGTACCTCTGTGGACCTGACATTCCTTACTACACCATACACCCCATCCCA 55581
Oy 168 ----- 168
Db 55580 GAAACCTCTGGATTAGGTCAAGACCTGCTTCATTTCCCTTGCGCTTAGGCGACTGGG 55521
Oy 168 ----- 168
Db 55520 GTGCTCTCACATTTCTGTGCAGCTTCAGAAATTTGTGCTCCCTTGACCTGAGATACA 55461
Oy 168 ----- 168
Db 55460 GGTGAGAGACAGTGGGAGTGTATTATCATTTATAGCATATATTTCTGAGATGTGGT 55401
Oy 168 ----- 168
Db 55400 TGGTGTGTTTGTCTCAGCTTGTCTGGGAATTTCTTCTGATCTGACATATACCTA 55341
Oy 169 ----- Aspleucysasphialaleu 175
Db 55340 TGAGAAATGACATATTCCTTAATGTTTGCATTTGCGAGATCTTGTGACATGCCCTG 55281
Oy 176 HISTSERHISAPGJaleu 182
Db 55280 CACATATGCGATGATGAGCTA 55260

RESULT 9
AC073896/c 182350 bp DNA linear HTG 24-AUG-2002
LOCUS Homo sapiens chromosome 12 clone RP11-977G19, WORKING DRAFT
DEFINITION AC073896
AC073896.37 GI:22450359
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULFILL; HTGS_ACTIVEFIN.
ORIGIN human.
Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 182350)
Munry,D.M., Adams,C., Adio-Oduola,B., All-oshan,F.R., Allen,C.,
Alstbrooks,S.L., Amaralungu,H.C., Arc,J.R., Ayale,M., Banks,T.,
Babalola,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Bulay,C., Burch,P., Burkelt,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Geo,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorelli,D.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,N., Holloway,C., Hollins,B.,
Homsli,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsen,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichteage,O., Liew,C., Liu,J., Liu,M., Louisged,H.,
Lozad,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Matlindala,A., Martinez,E.,
Massey,E., Mawhney,E., Mcleod,M.P., Meador,M., Mel,C., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabhat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newson,J., Newton,N., Nguyen,M., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Ogund,M., Okunodu,G.,
Oragunye,N., Oyler,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Qulles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I.,

```

Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Sytek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Umanal,K., Vasquez,L., Vera,V., Villalón,D., Vinsón,R., Wang,O.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleciyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.  
Direct Submission  
Unpublished  
2 (bases 1 to 182350)  
Worley,K.C.  
Submitted (01-JUL-2000) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 182350)  
Worley,K.C.  
Direct Submission  
Submitted (24-AUG-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Aug 23, 2002 this sequence version replaced gi:22296903.

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## Alignment Scores:

BASE COUNT 46453 a 45416 c 43855 g 46174 t 452 others

## FEATURES

## SOURCE

## SOURCE

## SOURCE

## SOURCE

## SOURCE

## SOURCE

## SOURCE

## SOURCE

## SOURCE

## SOURCE

## SOURCE

NOTE: Estimated insert size may differ from sequence length  
(see <http://www.hgsc.bcm.tmc.edu/docs/genbank.draft.data.html>).  
NOTE: This is a 'working draft' sequence. It currently  
consists of 5 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.  
1 2893: contig of 2893 bp in length  
2894 2993: gap of unknown length  
2994 6686: contig of 3693 bp in length  
6687 6787: gap of unknown length  
6787 17251: contig of 10465 bp in length  
17252 17352: gap of unknown length  
17352 85894: contig of 68543 bp in length  
85894 85995: gap of unknown length  
85995 182350: contig of 96356 bp in length.  
Location/Qualifiers  
1. 182350  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/clone="RP11-977G19"

```

Pred. No.: 1.61e-36 Length: 182350
Score: 481.00 Matches: 103
Percent Similarity: 63.80% Conservative: 1
Best Local Similarity: 63.19% Mismatches: 6
Query Match: 50.10% Indels: 53
DB: 2 Gaps: 1
US-10-082-502-17 (1-182) x AC073896 (1-182350)

QY 69 ValProTyAlaIarSergIuaIaHIsLeuThglUleuIeugluIleCysAspArg 88
|||||
Db 172027 GTCCTTATGCCCCGCTCAGAGGCCACCTCAGAGCTGCTGAGAGATATGTGACCG 171968
|||||
QY 89 MetLysGluTyGlyGluGluInIleAspProSerThrHisArgLysAntyValArgVal 108
|||||
Db 171967 ATGAGAGATATGAGGAGACAGATGATCTTCACCCATCGCAAGACTGACGTACGTGTA 171908
|||||
QY 109 ValGlyArgAsnGlyLysSerSergIuleuAspLeuGluGlyIleArgLysAspSerAsp 128
|||||
Db 171907 GTGGCCCGGAGAGAGAAATCCAGTGAACCTGACCAAGCATCCGAATCGACTGAGAT 171848
|||||
QY 129 IleSerGlyThrLeuLysPheAla----- 136
|||||
Db 171847 ATTAGGGGACCTCTCAAGTTGGGTGGAGCTATGGGAAATCGTACGTCTTTCGACAT 171788
|||||
QY 136 ----- 136
Db 171787 TAAGGGTTGTGGAGAACCAATTAGAAAGTATGCTGATCTTCTCTCTTGG 171728
|||||
QY 136 ----- 136
Db 171727 AAGTGAAGCGACAGACCTCTTATGTCCTGTGCACCCATTTCTCCCTGGTGGCAG 171668
|||||
QY 137 CysGluSerIleValGluGlyArgLysPheGluLeuIleGluPhePheSerArgGluAla 156
|||||
Db 171667 TGTGACAGCATTTGTGAGAGATACGAGATGACATCTTAACTCTTTCCAGAGACCT 171608
|||||
QY 157 AspAsnValLysAspLysLeuCysSerLysArgThrAsp-LeuCysAspHisAlaLeuH1 176
|||||
Db 171607 GACAAATGTTAAAGCAAACTTTCAGTAAAGCAAGTAAAGTCCGCCACCTTATCTC 171548
|||||
QY 176 sIleSer 178
|||||
Db 171547 CTCTTCT 171541

RESULT 10
AC023500 184762 bp DNA linear HMG 24-AUG-2002
LOCUS AC023500
DEFINITION Homo sapiens chromosome 12 clone RP11-183H16, WORKING DRAFT
ACCESSION AC023500
VERSION AC023500.32 GI:22095070
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 184762)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alshrocks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbarta,J., Benton,J., Bimarge,K., Blankenburg,K., Bonnin,D.,
Bouck,U., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burck,P., Burkett,C., Butrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homsli,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolyet,S., Joudah,S.,
Karlsom,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kraovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Liu,C., Liu,J., Liu,W., Louised,H.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Merindale,A., Martinez,E.,
Massey,E., Mayhew,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,K., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokkenko,S., Ogih,M., Okunolu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,D., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivera,M., Rojas,A., Rojiboken,I., Rolfe,M., Ruiz,S., Saverly,G.,
Scherer,S., Scott,G., Shen,H., Shoshart,N., Slason,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,O.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
Wu,C., Wu,X., Wu,Y., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
Submitted (15-FEB-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 184762)
Worley,K.C.
Direct Submission
Submitted (24-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Aug 4, 2002 this sequence version replaced gi:22094227.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Drafting Center Code: BCM
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HAPC
Center clone name: RP11-183H16
----- Summary Statistics
Sequencing vector: Plasmid;
Sequencing vector: M13;
Chemistry: Dye-primer Bodydy: 2% of reads
Chemistry: Dye-terminator Big Dye: 98% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 183845 bases at least Q40
Consensus quality: 184849 bases at least Q30
Consensus quality: 185447 bases at least Q20
Estimated insert size: 159026; sum-of-contrasts estimation
Quality coverage: 8x in Q20 bases; sum-of-contrasts estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 9 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
* 1 2040: contig of 2040 bp in length
* 2041 2140: gap of unknown length

```

```
*      2141      4170: contig of 2030 bp in length
*      4171      4270: gap of unknown length
*      4271      8294: contig of 4024 bp in length
*      8295      8394: gap of unknown length
*      8395      14196: contig of 5802 bp in length
*      14197      14296: gap of unknown length
*      14297      33076: contig of 18780 bp in length
*      33077      33176: gap of unknown length
*      33177      52295: contig of 19119 bp in length
*      52296      52395: gap of unknown length
*      52396      84185: contig of 31790 bp in length
*      84186      84285: gap of unknown length
*      84286      120023: contig of 35738 bp in length
*      120024      120123: gap of unknown length
*      120124      184762: contig of 64639 bp in length.
*      Location/Qualifiers
*      source          1..184762
                        /organism="Mus musculus"
                        /db_xref="taxon:9606"
                        /clone="RP11-183H16"

BASE COUNT   47425 a 44751 c 45101 g 46673 t      812 others
ORIGIN
Alignment Scores:
Pred. No.:      1.63e-36      Length:      184762
Score:          481.00      Matches:      103
Percent Similarity: 63.80%      Conservative: 6
Best Local Similarity: 63.19%      Mismatches: 1
Query Match:    50.10%      Indels:      53
DB:             2      Gaps:      1

US-10-082-502-17 (1-182) x AC033500 (1-184762)
QY      69      ValProTyrAlaArgSerGluAlaHisLeuThrGluLeuGluGluIleCysAspArg 88
         |||||||
Db      43486      GTGCGTTATGCCCGCTCAGAGAGCCCACTCACAGAGCTGCTGGAGAGATATGACCGG 43545
         |||||||
QY      89      MetLysGluTyrGluGlnIleAspProSerThrHisArgLysAsnTyrValArgVal 108
         |||||||
Db      43546      ATGAAGGAGTATGGGAGAACATTCATCTCCACCCATCCAGAACACTACGACGCTGA 43605
         |||||||
QY      109      ValGlyArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 128
         |||||||
Db      43606      GTGGCCGGGAATGAGAAATCCAGTGAACCTGACCAAGGCATCCGATGACATCAGAT 43665
         |||||||
QY      129      IleSerGlyThrLeuLysPheAla----- 136
         |||||||
Db      43666      ATTAGCGGCACCCCTCAATTGGCGGTGAGCTATGAGGATCGGTAGCTGTTCTTGACAT 43725
         |||||||
QY      136      ----- 136
         |||||||
Db      43726      TAAAGGGTTGTGGAGAACCAATTAGAAAGTATAGCGTATCTTGTCTCTCTTTGG 43785
         |||||||
QY      136      ----- 136
         |||||||
Db      43786      AAGTGAAGCGCAAGAGCGCTTATATGGCCCTGTGCACCCATTCTCCCTGGGTTGGCAG 43845
         |||||||
QY      137      CysGluSerIleValGluGluTyrGluAspLeuIleGluPhePheSerArgGluAla 156
         |||||||
Db      43846      TGTGAGACACTTGTGGAGCAATGACAGAGATGACATCAATTTCTTTCCCGAGAGCT 43905
         |||||||
QY      157      AspaAsnValLysAspLysLeuCysSerLysArgThrAsp-LeuCysAspHisAlaLeuH1 176
         |||||||
Db      43906      GACATATGTTAAAGCAAACTTGGAGTAAAGCAACAGTAGAGTCCGCCACTTATATCTC 43965
         |||||||
QY      176      sileser 178
         |||
Db      43966      CTGTCT 43972
         |||

RESULT 11
AC090489      AC090489      207424 bp      DNA      linear      ROD 26-JAN-2002
```

```
DEFINITION      Genomic sequence for Mus musculus, clone RP23-104010, complete
sequence.
ACCESSION      AC090489
VERSION      AC090489.8      GI:18376849
KEYWORDS
SOURCE
ORGANISM      Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS      McCombie W.R., de la Bastide M., Spiegel L., Preston R.,
Kirchoff K., Kuhl K., Nascimento L., Zutterer T., Balija V.,
Bell M., Baker J., Santos L., Miller B., Katzenberger F.,
Muller S., King L., Yang C., Palmer L., O'Shaughnessy A. and
Dedhia N.
TITLE      Genomic sequence for Mus musculus, clone RP23-104010, complete
sequence
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 207424)
AUTHORS      McCombie, W.R.
TITLE      Direct Submision
JOURNAL      Submitted (24-FEB-2001) Lila Annenberg Hazen Genome Sequencing
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
Harbor, NY 11724, USA
REFERENCE      3 (bases 1 to 207424)
AUTHORS      McCombie, W.R.
TITLE      Direct Submision
JOURNAL      Submitted (26-JAN-2002) Lila Annenberg Hazen Genome Sequencing
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
Harbor, NY 11724, USA
COMMENT
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest.
FEATURES
source          1..207424
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /clone="RP23-104010"
                /clone_11b="RPC1-23"
                48132..48165
                /note="We believe the assembly to be correct. The sequence
                is a mononucleotide (T) repeat in which the exact number
                of Ts is unknown. Other subclones in the region show one
                fewer T than that represented in the assembly."
misc_feature
BASE COUNT      52613 a 48671 c 49663 g 56477 t
ORIGIN
Alignment Scores:
Pred. No.:      5.61e-33      Length:      207424
Score:          446.50      Matches:      98
Percent Similarity: 48.77%      Conservative: 1
Best Local Similarity: 48.28%      Mismatches: 1
Query Match:    46.51%      Indels:      103
DB:             10      Gaps:      1

US-10-082-502-17 (1-182) x AC090489 (1-207424)
QY      69      ValProTyrAlaArgSerGluAlaHisLeuThrGluLeuGluGluIleCysAspArg 88
         |||||||
Db      75789      GTACCTTATGCCCGCTCAGAGAGCCCACTCACAGAGTTGCTTGAAGAGGTGTGTACCGA 75848
         |||||||
QY      89      MetLysGluTyrGluGlnIleAspProSerThrHisArgLysAsnTyrValArgVal 108
         |||||||
Db      75849      ATGAAGGAGTACGGGGAACAGATTGACCTTACCCACCGAGAACACTACGACGCTC 75908
         |||||||
QY      109      ValGlyArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 128
         |||||||
Db      75909      GTGACCGCGGAATGAGAAATCCAGTGAACCTTACAGGCGATCCGAAATTGACTCAGAT 75968
         |||||||
```



14201	15440	contig of 1640 bp in length
15841	15940	gap of unknown length
15941	17053	contig of 1113 bp in length
17054	17153	gap of unknown length
17154	18836	contig of 1683 bp in length
18837	18936	gap of unknown length
18937	20137	contig of 1201 bp in length
20138	20137	gap of unknown length
20238	21736	contig of 1499 bp in length
21737	21836	gap of unknown length
21837	23619	contig of 1783 bp in length
23619	23719	gap of unknown length
23720	25007	contig of 2188 bp in length
25008	26007	gap of unknown length
26008	27406	contig of 1399 bp in length
27407	27506	gap of unknown length
27507	28637	contig of 1331 bp in length
28638	28937	gap of unknown length
28938	30229	contig of 1292 bp in length
30230	30329	gap of unknown length
30330	32012	contig of 1683 bp in length
32013	32112	gap of unknown length
32113	33198	contig of 1086 bp in length
33199	33298	gap of unknown length
33299	35219	contig of 2221 bp in length
35220	35619	gap of unknown length
35620	37677	contig of 2058 bp in length
37678	37777	gap of unknown length
37778	38936	contig of 1159 bp in length
38937	39036	gap of unknown length
39037	40273	contig of 1237 bp in length
40274	40373	gap of unknown length
40374	41906	contig of 1533 bp in length
41907	42006	gap of unknown length
42007	44506	contig of 2500 bp in length
44507	44606	gap of unknown length
44607	46803	contig of 2197 bp in length
46804	46904	gap of unknown length
46904	49226	contig of 2323 bp in length
49227	49326	gap of unknown length
49327	55713	contig of 2107 bp in length
55714	55873	gap of unknown length
55874	58873	contig of 2060 bp in length
58874	61527	gap of unknown length
61528	61627	contig of 2654 bp in length
61628	63192	gap of unknown length
63193	65292	contig of 1565 bp in length
65293	65323	gap of unknown length
65324	65362	contig of 2070 bp in length
65363	65462	gap of unknown length
65463	67438	contig of 1976 bp in length
67439	67538	gap of unknown length
67539	70505	contig of 2967 bp in length
70506	70605	gap of unknown length
70606	74524	contig of 3919 bp in length
74525	74624	gap of unknown length
74625	76737	contig of 2113 bp in length
76738	76837	gap of unknown length
76838	80133	contig of 3296 bp in length
80134	80233	gap of unknown length
80234	83122	contig of 2889 bp in length
83123	83222	gap of unknown length
83223	85731	contig of 3509 bp in length
85732	86831	gap of unknown length
86832	89929	contig of 3098 bp in length
89930	90029	gap of unknown length
90030	93107	contig of 3078 bp in length
93108	93207	gap of unknown length
93208	97419	contig of 4212 bp in length

FEATURES	97420	97519: gap of unknown length
source	101666	Location/Qualifiers
	1. 101666	/organism="Rattus norvegicus"
		/db_xref="taxon:10116"
		/clone="CH230-319B6"
BASE COUNT	24650 a 22934 c 23569 g 23059 t	7454 others
ORIGIN		
Alignment Scores:		
Pred. No.:	2.64e-32	Length: 101666
Score:	436.00	Matches: 99
Percent Similarity:	49.75%	Conservative: 0
Best Local Similarity:	49.75%	Mismatches: 1
Query Match:	45.42%	Indels: 99
DB:	2	Gaps: 1
US-10-082-502-17 (1-182) x AC109891 (1-101666)		
Qy	69	ValProtyAlaArgSerGluAlaHisLeuThrGluLeuGluGluIleCysAspArg 88
Db	97572	GTTCCCTTATGCCCGCTCAGAGGCCACCTCAGCGAGTTGCTTGAGGAAATTTGACCGA 97631
Qy	89	MettysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValArgVal 108
Db	97632	ATGAAGAGATATGGGGAACAGATGATGACCTTCTTACCCAGCCGAAAGAACTAGCTCCGTGT 97691
Qy	109	ValGlyArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 128
Db	97692	GTCGGCCGGGACGGAGGATCCAGTAACTAGACTTACAGGATATCCGAAATGATTCAGAT 97751
Qy	129	IleSerGlyThrLeuLysPhe----- 135
Db	97752	ATCAAGTGGACCCCTCAAGTTTCCGGTGAATTAAGCTGTGCGGTAGTACCTGGTTCTTG 97811
Qy	135	----- 135
Db	97812	GGAGTTCAATAGGAAATCTCGGCGACAGATTTGGAGATCGCGCGGATATGCTTTCATG 97871
Qy	135	----- 135
Db	97872	CTTTGTAACCGGGCAGTGGAGAGCAGTGTGGAAATTTGGCACTAGTCACTAGT 97931
Qy	135	----- 135
Db	97932	TCTAGGCCACGTGGAGCGCATATCAATAAACCTGTCTTAATAATCCCACTGCCACCAA 97991
Qy	135	----- 135
Db	97992	AAACACGAGAGAGACTGAGCGCGTCTCTGCCCCGCGTGAAGTCTGGGTCACTTCT 98051
Qy	136	----- Ala-CysGluSerIleValGluGluTyrGluAspGluLeuIleG1 150
Db	98052	CTGTCTTCTTGGTCGCGCAGTGTGAAGAGATTTGGAGAGATGACGAGTGAATTAATTGA 98111
Qy	150	uphePheSerArgGluAlaAspAsnValLysAspLysLeuGlySerLysArgThr 168
Db	98112	ATTCTTTTCAAGAGAGCGCTGACACAGCTTAAGACAACTTTGCAGTAAGCCGACA 98166
RESULT 13		
LOCUS	AF186113	649 bp mRNA linear PRI 13-JAN-2000
DEFINITION	Homo sapiens putative secreted protein ZS1G9 (ZS1G9) mRNA, complete cds.	
ACCESSION	AF186113	
VERSION	AF186113.1	GI:6014631
KEYWORDS		
SOURCE	Homo sapiens.	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
REFERENCE	1 (bases 1 to 649)	

**AUTHORS** Sheppard, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J. and O'Hara, P.

**TITLE** Homo sapiens putative secreted protein

**JOURNAL** Unpublished

**REFERENCE** 2 (bases 1 to 649)

**AUTHORS** Sheppard, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J. and O'Hara, P.

**TITLE** Direct Submission

**JOURNAL** Submitted (14-SEP-1999) Biomedical Informatics, ZymoGenetics, Inc., 1201 Eastlake Ave., East, Seattle, WA 98102, USA

**FEATURES**

**source**

1..649

/organism="Homo sapiens"

/db\_xref="dbest:R15306"

/db\_xref="taxon:9606"

/chromosome="12"

/map="12q15"

/clone="IMAGE:29675"

1..649

/gene="ZS1G9"

104..358

/gene="ZS1G9"

/codon\_start=1

/product="putative secreted protein ZS1G9"

/protein\_id="AA01431.1"

/db\_xref="GI:6014632"

/translation="MKCGWMLALLGALLCTANARRSODLHGCACRALVDELEWELIAQVDPKRTOMGSFRINDGSQSVVEVTVTPPNKVAHSFG"

**BASE COUNT** 175 a 150 c 171 g 149 t 4 others

**ORIGIN**

Alignment Scores:

Pred. No.: 2.76e-28 Length: 649

Score: 369.00 Matches: 69

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 38.44% Indels: 0

Gaps: 0

US-10-082-502-17 (1-182) x AF186113 (1-649)

OY 1 MetlysglyttrpglyttrpneuallaleuLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20

DB 104 ATGAAGGCTGGGGTGGCTGGCCCTCTCTGGGGGCGCTGGAGACCGCTGGGCT 163

OY 21 ArgArgserGlnAspLeuHiscysGlyAlaCysArgAlaLeuValAspGlyLeuGlyTrp 40

DB 164 CGAGGAGCCAGGATCTCCACTGTGGAGCATGCGAGGCTCTGTGGATGAACAGATGG 223

OY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60

DB 224 GAAATGGCCAGGCTGGAGCCCAAGAACACCATTCAGATGGAGATCTTCCGATCAATCCA 283

OY 61 AspGlySerGlnSerValValGluVal 69

DB 284 GATGGCAGCCAGTCAGTGTGGAGGTA 310

**RESULT 14**

BC001027

LOCUS BC001027 824 bp mRNA linear PRI 12-JUL-2001

DEFINITION Homo sapiens, transmembrane protein 4, clone MGC:1545

IMAGE:3344786, mRNA, complete cds.

ACCESSION BC001027.1 GI:12654402

KEYWORDS MGC.

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 824)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (17-NOV-2000) National Institutes of Health, Mammalian

**REMARK**

**COMMENT**

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILN)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)

Gallthersburg, Maryland.

Web site: <http://www.nisc.nih.gov/>

Contact: [nisc\\_mgcenhgri.nih.gov](mailto:nisc_mgcenhgri.nih.gov)

Shenchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McElroy, J.C., McDowell, J., Pearson, R., Snyder, B., Stantipop, S., Thomas, P.J., Tlonsen, E.E., Touchman, D.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Zhang, L., H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN at: <http://image.llnl.gov>

Series: IRAL Plate: 5 Row: P Column: 13

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6014631.

**FEATURES**

**source**

1..824

/organism="Homo sapiens"

/db\_xref="locusid:10330"

/db\_xref="taxon:9606"

/clone="MGC:1545 IMAGE:3344786"

/rissue\_type="Eye, retinoblastoma"

/clone\_1lb="NIH\_MGC\_16"

/lab\_host="DH10B-R"

/note="Vector: pOTB7"

250..504

/codon\_start=1

/product="transmembrane protein 4"

/protein\_id="AA01027.1"

/db\_xref="GI:12654403"

/translation="MKCGWMLALLGALLCTANARRSODLHGCACRALVDELEWELIAQVDPKRTOMGSFRINDGSQSVVEVTVTPPNKVAHSFG"

**BASE COUNT** 221 a 194 c 233 g 176 t

**ORIGIN**

Alignment Scores:

Pred. No.: 3.71e-28 Length: 824

Score: 369.00 Matches: 69

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 38.44% Indels: 0

Gaps: 0

US-10-082-502-17 (1-182) x BC001027 (1-824)

OY 1 MetlysglyttrpglyttrpneuallaleuLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20

DB 250 ATGAAGGCTGGGGTGGCTGGCCCTCTCTGGGGGCGCTGGAGACCGCTGGGCT 309

OY 21 ArgArgserGlnAspLeuHiscysGlyAlaCysArgAlaLeuValAspGlyLeuGlyTrp 40

DB 310 CGAGGAGCCAGGATCTCCACTGTGGAGCATGCGAGGCTCTGTGGATGAACATGAG 369

OY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60

DB 370 GAAATGGCCAGGCTGGAGCCCAAGAACACCATTCAGATGGAGATCTTCCGATCAATCCA 429

OY 61 AspGlySerGlnSerValValGluVal 69

DB 430 GATGGCAGCCAGTCAGTGTGGAGGTA 456

RESULT 15  
AC103156/c  
LOCUS  
DEFINITION  
AC103156  
AC103156 3  
AC103156.3  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE

AUTHORS

69753 bp DNA linear HTG 13-JUL-2002  
Rattus norvegicus clone CH230-208H17, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 42 unordered pieces.  
AC103156  
AC103156.3  
HTG: HTGS, PHASE1.  
Norway rat.  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

COMMENT

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 69753)  
Muzny, D.M., Adams, C., Adlo-Oduola, B., All-otman, F.R., Allen, C.,  
Albrooks, S.L., Amaral, H.C., Are, J.R., Ayala, M., Banks, T.,  
Barbata, J., Benton, J., Blum, K., Blankenburg, K., Bonin, D.,  
Bouch, J., Bowler, S., Brilleman, M., Brown, E., Brown, M., Bryant, N.P.,  
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
DeLaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Din, H.H.,  
Douthett, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Einhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabis, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hayes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hughes, M., Holloway, C., Hollins, B.,  
Hosni, P., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kretovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtenberg, O., Liu, C., Liu, J., Liu, W., Louisgeed, H.,  
Lorado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mathew, E., McLeod, M.P., Meador, M., Mei, G., Melzer, M.,  
Miner, G., Miner, Z., Mitchell, T., Monabhat, R., Morgan, M., Morris, S.,  
Moore, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokenko, S., Oguh, M., Okunnu, G.,  
Orduyev, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peterson, L., Pickens, R., Primus, E., Pu, L.L., Qules, M., Ren, Y.,  
Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Severy, G.,  
Scheier, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I.,  
Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Sytek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Umanal, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G. and Gibbs, R.

Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GJIX  
Center clone name: CH230-208H17  
----- Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 15506 bases at least Q40  
Consensus quality: 17125 bases at least Q30  
Consensus quality: 18271 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a "working draft" sequence. It currently  
\* consists of 42 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 1366: contig of 1366 bp in length  
\* 1467: gap of unknown length  
\* 2632: contig of 1166 bp in length  
\* 2733: gap of unknown length  
\* 3919: contig of 1187 bp in length  
\* 4019: gap of unknown length  
\* 4199: contig of 1150 bp in length  
\* 5269: gap of unknown length  
\* 6813: contig of 1344 bp in length  
\* 6913: gap of unknown length  
\* 8119: contig of 1206 bp in length  
\* 8219: gap of unknown length  
\* 9384: contig of 1165 bp in length  
\* 9484: gap of unknown length  
\* 10996: contig of 1512 bp in length  
\* 11096: gap of unknown length  
\* 12554: contig of 1458 bp in length  
\* 12654: gap of unknown length  
\* 14376: contig of 1722 bp in length  
\* 14476: gap of unknown length  
\* 15790: contig of 1314 bp in length  
\* 15890: gap of unknown length  
\* 15991: gap of unknown length  
\* 17010: contig of 1120 bp in length  
\* 17110: gap of unknown length  
\* 18308: contig of 1198 bp in length  
\* 18408: gap of unknown length  
\* 19786: contig of 1378 bp in length  
\* 19886: gap of unknown length  
\* 21145: contig of 1259 bp in length  
\* 21245: gap of unknown length  
\* 22539: contig of 1294 bp in length  
\* 22639: gap of unknown length  
\* 23861: contig of 1222 bp in length  
\* 23961: gap of unknown length  
\* 25395: contig of 1434 bp in length  
\* 25495: gap of unknown length  
\* 26515: contig of 1020 bp in length  
\* 26615: gap of unknown length  
\* 28329: contig of 1714 bp in length  
\* 28429: gap of unknown length  
\* 30046: contig of 1617 bp in length  
\* 30146: gap of unknown length  
\* 31613: contig of 1467 bp in length  
\* 31713: gap of unknown length  
\* 33120: contig of 1407 bp in length  
\* 33220: gap of unknown length  
\* 34615: contig of 1395 bp in length  
\* 34715: gap of unknown length  
\* 35867: contig of 1152 bp in length  
\* 35967: gap of unknown length  
\* 37671: contig of 1704 bp in length



37672 37771: gap of unknown length  
37772 39476: contig of 1705 bp in length  
39477 39576: gap of unknown length  
41221 41221: contig of 1645 bp in length  
41222 41321: gap of unknown length  
41322 41330: contig of 1809 bp in length  
41331 43231: gap of unknown length  
43231 44734: contig of 1504 bp in length  
44735 44834: gap of unknown length  
44835 46199: contig of 1365 bp in length  
46200 46299: gap of unknown length  
46300 49061: contig of 2762 bp in length  
49062 49161: gap of unknown length  
49162 51059: contig of 1898 bp in length  
51060 51159: gap of unknown length  
51160 53067: contig of 1908 bp in length  
53068 53167: gap of unknown length  
53168 54641: contig of 1474 bp in length  
54642 54741: gap of unknown length  
54742 57496: contig of 2755 bp in length  
57497 57596: gap of unknown length  
57597 59003: contig of 1407 bp in length  
59004 59103: gap of unknown length  
59104 60621: contig of 1518 bp in length  
60622 60721: gap of unknown length  
60722 62709: contig of 1988 bp in length  
62710 62809: gap of unknown length  
62810 64884: contig of 2075 bp in length  
64885 64984: gap of unknown length  
64985 67653: contig of 2669 bp in length  
67654 67753: gap of unknown length  
67754 69753: contig of 2000 bp in length.

## FEATURES

1..69753  
Location/Qualifiers  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-208H17"

BASE COUNT 15260 a 17000 c 17392 g 15839 t 4262 others  
ORIGIN

## Alignment Scores:

Pred. No.: 1..1e-21  
Score: 328.50  
Percent Similarity: 56.67%  
Best Local Similarity: 55.83%  
Query Match: 34.22%  
DB: 2  
Caps: 1

US-10-082-502-17 (1-182) x AC103156 (1-69753)

OY 1 MetysGLYTPGGLYTPLEUAlaLeuLeuGLyAlaLeuLeuGLYThraLATrPaLa 20  
|||||  
Db 15128 ATGAAGGCTGGGGTGGCTGGCCCTACTTTTGGGGCTCTGCTGGGAACCTAGCGGCT 15069  
OY 21 ATGATsErgInAsPLeuHIsCySGLy----- 29  
|||||  
Db 15068 CGAAGGAGCCAGGATCTACACTGTGGAGTAAAGCAAAATACAGTGAAGAAGTGGGA 15009  
OY 29 ----- 29  
Db 15008 GGAAGCCGACCTTGGAGTCATTAGATCCAAAGCTTGGAAACAAGATGATGAACCC 14949  
OY 30 -----AlaCys 31  
Db 14948 TTGGGTTGACCCATCCCTCAGATCCACATCCCTATTCTCTTCTGGCCGACGCTGC 14889  
OY 32 ATGAlaLeuValAsPGLuLeuGLuIleAlaGlnValAsPProLySLyThrlle 51  
|||||  
Db 14888 AGGGCTCTGGTGAATTAAGTGGGAATGGCCCGCTGACCCCAAGAAGACCAATT 14829  
OY 52 GlnMetGLySerPheArgIleAsnProAsPGLySerGlnSerValValGluValProTyr 71  
|||||  
Db 14828 CAGATGGATCTTCGGAATCAATCCAGATGGACGACGACGATGAGTGTGGAGGTAACTTAC 14769

Search completed: March 31, 2003, 03:37:06  
Job time : 1882.3 secs



score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

(million alignments)  
10796.945 Million cell updates/sec

Title: US-10-082-502-18  
Perfect score: 1069  
Servanno: 1222-74

**SECRET**

Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 4109280

Maximum DB seq length: 200000000000

Post-processing: Minimum Match 08

Listing first 45 summaries

**Palcauabé**

```

1:  gb_ba:*
2:  gb_htg:*
3:  gd_in:*
4:  gb_com:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pt:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_com:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htg_hum:*
40: em_htg_mus:*
41: em_htggo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

C	No.	Score	Query	Match	Length	DB	ID	Description
1	1017	95.1	1017	10	AF186115		AF186115 Mus musc	
2	795.2	74.4	158198	2	AC012013		AC012013 Homo sapi	
3	794.8	74.3	1098	10	BC008261		BC008261 Homo sapi	
4	691	64.6	1080	9	AY032624		AY032624 Mus musc	
5	691	64.6	1210	6	AX464040		AX464040 Homo sapi	
6	606.6	56.7	814	6	AB440456		AB440456 Sequence	
7	606.6	56.7	814	6	AB015631		AB015631 Homo sapi	
8	321	30.0	207424	10	AC090489		AC090489 Genomic f	
9	307	28.7	824	9	BC001027		BC001027 Homo sapi	
10	248.6	23.1	101666	2	AC109891		AC109891 Rattus no	
11	246.8	23.1	649	9	AF186113		AF186113 Homo sapi	
12	221	20.7	564	6	AX440375		AX440375 Sequence	
13	177.4	16.6	155023	2	AC025574		AC025574 Homo sapi	
14	177.4	16.6	182350	2	AC073896		AC073896 Homo sapi	
15	177.4	16.6	184762	2	AC023500		AC023500 Homo sapi	
16	126.4	11.8	69753	2	AC103156		AC103156 Rattus no	
17	88.2	8.3	2378	9	AC094445		AC094445 Homo sapi	
18	87.2	8.2	95437	2	AC096788		AC096788 Rattus no	
19	66.8	6.2	147	6	AX333297		AX333297 Sequence	
20	66.8	6.2	147	6	AX408688		AX408688 Sequence	
21	56.2	5.3	622	5	AB055671		AB055671 Dnlo rer	
22	52.2	4.9	125020	6	AX429315		AX429315 Sequence	
23	51.6	4.8	126	6	AX333296		AX333296 Homo sapi	
24	51.6	4.8	126	6	AX408687		AX408687 Sequence	
25	50	4.7	105565	6	AC104103		AC104103 Mus muscu	
26	49.4	4.6	210631	2	AC129184		AC129184 Mus muscu	
27	49.4	4.6	224840	2	AC027284		AC027284 Mus muscu	
28	49	4.6	2269	10	AF220121		AF220121 Mus muscu	
29	48.6	4.5	168418	9	AC009403		AC009403 Homo sapi	
30	48.4	4.5	218900	2	AC125202		AC125202 Mus muscu	
31	48.2	4.5	2136	9	HSMB04456		HSMB04456 Homo sapi	
32	47.8	4.5	191793	2	AF001897		AF001897 Homo sapi	
33	47.6	4.5	1641	9	BC011444		BC011444 Homo sapi	
34	47.6	4.5	1651	9	BC003108		BC003108 Homo sapi	
35	47.6	4.5	1651	9	HSMB03247		HSMB03247 Homo sapi	
36	47.6	4.5	165566	9	AC008060		AC008060 Homo sapi	
37	47.4	4.4	2326	9	BC011860		BC011860 Homo sapi	
38	47.4	4.4	2433	3	AF007768		AF007768 Choriiston	
39	47.2	4.4	1583	10	BC034896		BC034896 Homo sapi	
40	47.2	4.4	1688	9	AK026028		AK026028 Mus muscu	
41	47	4.4	2870	9	HSMB00882		HSMB00882 Homo sapi	
42	47	4.4	167865	2	AC033471		AC033471 Homo sapi	
43	46.8	4.4	1260	6	AX380553		AX380553 Sequence	
44	46.8	4.4	1630	9	BC028028		BC028028 Homo sapi	
45	46.8	4.4	169794	2	AC004688		AC004688 Plasmodiu	

## ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
AF186115	AF186115	1017 bp mRNA Mus musculus putative secreted protein ZS1g9 (Zs1g9) mRNA, complete cds.	AF186115	AF186115	GI:6014635	Mus musculus. Mus musculus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathu; Muridae; Murinae; Mus.	Sheppard,P., Jelinek,L., Whitmore,T., Blumberg,H., Lehner,J. and O'Hara,P.	

TITLE Mus musculus putative secreted protein  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1017)  
AUTHORS Sheppard, P., Jelinek, L., Whitmore, T., Blumberg, H., Lehnert, J. and O'Hara, P.  
TITLE Direct Submission  
JOURNAL Submitted (14-SEP-1999) Biomolecular Informatics, ZymoGenetics, Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA  
FEATURES  
source  
1..1017  
/organism="Mus musculus"  
/db\_xref="dbEST:AA049839"  
/db\_xref="taxon:10090"  
/clone="IMAGE:476061"  
1..1017  
/gene="Zs19"  
/gene="Zs19"  
343..891  
/gene="Zs19"  
/product="putative secreted protein ZS19"  
/protein\_id="AA01433.1"  
/db\_xref="GI:6014636"  
translation="MKGMWLLALLGLVLTAMARRSODLHGACRALVDELEWETAR  
VDPKTIOMGSEFRINPDGOSVVEPYARSPAHJTELEFEVCDMEKEGEGIDPSIHR  
KNYRVYSRNESESELDIGIRIDSDISGTLKRFACSEIVEDEBELIEFFSRADNPK  
DKLSKRTDLCDHLHSHDEL"  
BASE COUNT 248 a 250 c 315 g 204 t  
ORIGIN  
Query Match 95.1%; Score 1017; DB 10; Length 1017;  
Best Local Similarity 100.0%; Pred. No. 6e-265;  
Matches 1017; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 16 GGGGCTCTGCTGCTCGGAGGCGCTCTCTAAAGCTGCTGCTGCGCGAGATTGGAGG 75  
DB 1 GGGGCTCTGCTGCTCGGAGGCGCTCTCTAAAGCTGCTGCTGCGCGAGATTGGAGG 60  
QY 76 GGGGCTCTGCTGCTCGGAGGCGCTCTCTAAAGCTGCTGCTGCGCGAGATTGGAGG 135  
DB 61 GGGGCTCTGCTGCTCGGAGGCGCTCTCTAAAGCTGCTGCTGCGCGAGATTGGAGG 120  
QY 136 GTGCTAGTTAAGTGGGGAAGCTGGGCGAGGCGTCACTGGCGACCCGTAAGCTGGGGG 195  
DB 121 GTGCTAGTTAAGTGGGGAAGCTGGGCGAGGCGTCACTGGCGACCCGTAAGCTGGGGG 180  
QY 196 GAGCGGAGCGCTCTGGAAGAGCCGCGAGAGCCCGCTTTTCCAGCCAGCTGTAGGGT 255  
DB 181 GAGCGGAGCGCTCTGGAAGAGCCGCGAGAGCCCGCTTTTCCAGCCAGCTGTAGGGT 240  
QY 256 TGGGACCCAGCAAAAAGTGAAGTCCGCTCTTCCAGACCTGGGCGACGCGG 315  
DB 241 TGGGACCCAGCAAAAAGTGAAGTCCGCTCTTCCAGACCTGGGCGACGCGG 300  
QY 316 CGGCGCTGGGAGCAGAGTGGAGCGACCTGTACTAAGATGAAGAGCTGGGGTGG 375  
DB 301 CGGCGCTGGGAGCAGAGTGGAGCGACCTGTACTAAGATGAAGAGCTGGGGTGG 360  
QY 376 CTAGCCCTACTTTTGGGGTCTGCTGTGGAACTGCTGGCTCGAAGAGCCCAAGATCTA 435  
DB 361 CTAGCCCTACTTTTGGGGTCTGCTGTGGAACTGCTGGCTCGAAGAGCCCAAGATCTA 420  
QY 436 CACTGTGAGCTTGCAGGGCTCTGCTGTGATGATTAAGTGGGAAATGGCCGCGTGAC 495  
DB 421 CACTGTGAGCTTGCAGGGCTCTGCTGTGATGATTAAGTGGGAAATGGCCGCGTGAC 480  
QY 496 CCCAAGAGACATTCAGATGGATCTTCGAATCATCCAGATGGAGCAGCAGCATTT 555  
DB 481 CCCAAGAGACATTCAGATGGATCTTCGAATCATCCAGATGGAGCAGCAGCATTT 540  
QY 556 GTGAGAGTACCTTATGCGCGCTCAGAGGCGCAGCCTCAGAGATTGCTTGAGAGGTGTGT 615  
DB 541 GTGAGAGTACCTTATGCGCGCTCAGAGGCGCAGCCTCAGAGATTGCTTGAGAGGTGTGT 600

QY 616 GACCGAATGAGAGTACGCGGAGACAGATTTGACCTTCTACCCACCGAAGACTACGTA 675  
DB 601 GACCGAATGAGAGTACGCGGAGACAGATTTGACCTTCTACCCACCGAAGACTACGTA 660  
QY 676 GCGGTGAGCGGGAATGAGAAATCCAGTGAATGACTTACAGGCAATCCGATTTGAC 735  
DB 661 GCGGTGAGCGGGAATGAGAAATCCAGTGAATGACTTACAGGCAATCCGATTTGAC 720  
QY 736 TCAGATATCAGCGGACCCCTCAAGTTTGGCTGAGAGCATTTGGAAGATACAGAT 795  
DB 721 TCAGATATCAGCGGACCCCTCAAGTTTGGCTGAGAGCATTTGGAAGATACAGAT 780  
QY 796 GAGCTTATGAAATTTCTTCCAGAGAGGCTGACAAAGTTAAAGACAACTTGCAGTAAG 855  
DB 781 GAGCTTATCAGAAATTTCTTCCAGAGAGGCTGACAAAGTTAAAGACAACTTGCAGTAAG 840  
QY 856 CGGACAGATCTATGACCATTCCTCGACAGATTCACAGATGACCTATGATCATCTGGA 915  
DB 841 CGGACAGATCTATGACCATTCCTCGACAGATTCACAGATGACCTATGATCATCTGGA 900  
QY 916 GCAAGCAGCTTACACCAAACTGATGGAACACCCCGAGAGGGAAGATGCGAGATTGC 975  
DB 901 GCAAGCAGCTTACACCAAACTGATGGAACACCCCGAGAGGGAAGATGCGAGATTGC 960  
QY 976 CTTTATATATACGTTTTTATGGAATGAACTGAAAAAACTTTGAAACCGAAGTA 1032  
DB 961 CTTTATATATACGTTTTTATGGAATGAACTGAAAAAACTTTGAAACCGAAGTA 1017  
RESULT 2  
AC012013  
LOCUS  
DEFINITION Homo sapiens chromosome 12 clone RP11-764L14, WORKING DRAFT  
SEQUENCE, 5 unordered pieces.  
AC012013  
VERSION  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
SOURCE  
Homo sapiens  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 158198)  
Muzny, D.M., Adams, C., Adlo-Oduola, B., Alt-osman, F.R., Allen, C.,  
Alsbrooks, S.L., Amaratunga, H.C., Are, J.R., Banks, T., Barbarta, J.,  
Benton, J., Bimarge, K., Blankenburg, K., Bonnin, D., Bouck, J.,  
Bowie, S., Brielev, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,  
Burck, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F.,  
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,  
Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C.,  
Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,  
Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,  
Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,  
Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C.,  
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,  
Foster, P., Frantz, P., Gabisl, A., Gao, J., Garcia, A., Garner, T.,  
Hollins, B., Homel, F., Howard, S., Huber, J., Huliy, S., Hume, J.,  
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,  
Joudan, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,  
Kovar, C., Kratovic, J., Kuresh, A., Landry, N., Leal, B., Lewis, L.C.,  
Lewis, L., Li, J., Li, J., Lichtharg, O., Lieu, C., Liu, J., Liu, W.,  
Louliseg, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,  
Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M.,  
Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabhat, K.,  
Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,  
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S.,  
Ogih, M., Okumou, G., Oragunye, N., Oyedelo, R., Pace, A., Payton, B.,  
Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,  
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojudoan, I., Rolfe, M.,  
Rulz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shooshitari, N.,



COMMENT Contact: MGC help desk  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
Contact: amg@bcm.tmc.edu  
Gunnarane, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Hale, S.M.,  
Yoon, V.S., Kowls, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LIML at: http://image.liml.gov  
Series: IRAX Plate: 5 Row: b Column: 23  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 9903606.  
Location/Qualifiers  
1. 798  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="129, C57BL/6J, FVB/N"  
/clone="MGC:6853 IMAGE:2650612"  
/tissue="MGC:6853 IMAGE:2650612"  
/tissue\_type="mammary tumor. Brcal-/fl: MMTV-Cre model. 10  
months old, gross tissue."  
/clone\_11b="NCI\_CGAP\_Mam3"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
115. 663  
/codon\_start=1  
/product="transmembrane protein 4"  
/protein\_id="AA08261.1"  
/db\_xref="GI:14198401"  
/db\_xref="LocusID:56530"  
/translation="MKMGWALLLGLGTLGTMARSRDILHCGACRALVDELEFIAR  
VDPKKTOMGFRINPDGOSQVVEPYPARSAHLTELLEECIDMKKEGQIDPSTHR  
KNYRVVGRNGESSELDLQIRIDSDISGLTKFACISIVEEYEDLEIEFSREADNVK  
DKLCSKRTDLCDAHLHRSDEL"

CDS  
BASE COUNT 227 a 182 c 228 g 161 t  
ORIGIN

Query Match 74.3%; Score 794.8; DB 10; Length 798;  
Best Local Similarity 99.7%; Pred. No. 1.2e-204;  
Matches 796; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 244 AGCTCTAGAGGTGGGACCCACAGAAACAAAGTGAAGTCCGGCTCTTCCAGAGCCT 303  
DB 1 ACCTGCTAGAGGTGGGACCCACAGAAACAAAGTGAAGTCCGGCTCTTCCAGAGCCT 60  
OY 304 GGGCCACGGCGGCGCGCTGGGAGAGAGAGTGGAGCAGCCCTGTTACATAAAGTGA 363  
DB 61 GGGCCACGGCGGCGCGCTGGGAGAGAGAGTGGAGCAGCCCTGTTACATAAAGTGA 120  
OY 364 GGGTGGGGTGGCTAGCCCTACTTTTGGGGCTCTGCTGGAGACGCTGGGCTGGAAG 423  
DB 121 GGGTGGGGTGGCTAGCCCTACTTTTGGGGCTCTGCTGGAGACGCTGGGCTGGAAG 180  
OY 424 AGCCAAAGATCTACATCTGAGACCTTGCAGAGGCTCTGTGGATGAATTAAGTGGGAAT 483  
DB 181 AGCCAAAGATCTACATCTGAGACCTTGCAGAGGCTCTGTGGATGAATTAAGTGGGAAT 240  
OY 484 GCCCGCGTGGAGCCCAAGAAAGCATTCAGATGGATCCTTCCGATCAATCCAGATGCG 543  
DB 241 GCCCGCGTGGAGCCCAAGAAAGCATTCAGATGGATCCTTCCGATCAATCCAGATGCG 300  
OY 544 AGCCAGTAGTTTGGAGAGTACCTTATGCGCGCTAGAGAGCCCACTCAGAGAGTTGCT 603  
DB 301 AGCCAGTAGTTTGGAGAGTACCTTATGCGCGCTAGAGAGCCCACTCAGAGAGTTGCT 360  
OY 604 GAGGAGGTGTGACCGGAATGAAGAGTACGGGGAACAGATTGACCTTTACCAACCGC 663  
DB 604 GAGGAGGTGTGACCGGAATGAAGAGTACGGGGAACAGATTGACCTTTACCAACCGC 663

DB 361 GAGGAGGTGTGACCGGAATGAAGAGTACGGGGAACAGATTGACCTTTACCAACCGC 420  
OY 664 AAGAACTACGATGCGGCTCGTGGAGCCGAATGGAATCCAGTGAATACCTAGAGGCG 723  
DB 421 AAGAACTACGATGCGGCTCGTGGAGCCGAATGGAATCCAGTGAATACCTAGAGGCG 480  
OY 724 ATCCGAATTTGACTCAGATATTCAGAGCGGACCCCTCAAGTTTGGCTGTGAGACATTTGGA 783  
DB 481 ATCCGAATTTGACTCAGATATTCAGAGCGGACCCCTCAAGTTTGGCTGTGAGACATTTGGA 540  
OY 784 GAATCGAGAGATGAGCTTTGCAATTTCTTCACAGAGAGCGTGAACCTTAAACACAA 843  
DB 541 GAATCGAGAGATGAGCTTTGCAATTTCTTCACAGAGAGCGTGAACCTTAAACACAA 600  
OY 844 CTTTGAGTAGAGCGGACAGATATTCAGTACATGCGCTGCACAGATTCACAGATGAGCT 903  
DB 601 CTTTGAGTAGAGCGGACAGATATTCAGTACATGCGCTGCACAGATTCACAGATGAGCT 660  
OY 904 TGAATCTCTGAGAGCAAGACGCTTACACCAACCTGATGGAACACCCCGAGAGGGAAGA 963  
DB 661 TGAATCTCTGAGAGCAAGACGCTTACACCAACCTGATGGAACACCCCGAGAGGGAAGA 720  
OY 964 TGGCAGATTTGCTTTTATATATACGTTTATGAAATGAACTGAAATACCTTTGAA 1023  
DB 721 TGGCAGATTTGCTTTTATATATACGTTTATGAAATGAACTGAAATACCTTTGAA 780  
OY 1024 CCGAAAGTAAAAA 1041  
DB 781 CCGAAAGTAAAAA 798

RESULT 4  
AY032624 1080 bp mRNA linear PRI 18-APR-2002  
LOCUS  
DEFINITION Homo sapiens saposin-like protein mRNA, complete cds.  
ACCESSION AY032624  
VERSION AY032624.1 GI:20196198  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1080)  
AUTHORS Bornhauser, B.C., Olsson, P.-A. and Lindholm, D.  
TITLE NSAP is a novel saposin-like protein that interacts with MIR and stimulates neurite outgrowth  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1080)  
AUTHORS Olsson, P.-A. and Lindholm, D.  
TITLE Direct Submission  
JOURNAL Submitted (13-APR-2001) Neuroscience, Uppsala University, Husergatan 3, Uppsala 75123, Sweden  
FEATURES  
source  
1. 1080  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/map="12q15"  
419. 967  
/note="NSAP: encodes a type 2 membrane signal and a putative ER retention motif"  
/product="saposin-like protein"  
/protein\_id="AAK38148.1"  
/db\_xref="GI:20196199"  
/translation="MKMGWALLLGLGTLGTMARSRDILHCGACRALVDELEFIAR  
VDPKKTOMGFRINPDGOSQVVEPYPARSAHLTELLEECIDMKKEGQIDPSTHR  
KNYRVVGRNGESSELDLQIRIDSDISGLTKFACISIVEEYEDLEIEFSREADNVK  
DKLCSKRTDLCDAHLHRSDEL"

BASE COUNT 238 a 269 c 343 g 230 t  
ORIGIN

Query Match 64.6%; Score 691; DB 9; Length 1080;

Best Local Similarity 84.4%; Pred. No. 1,9e-176;  
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

```

OY 12 GAGGGGGGCTCGCTCGGAGCGCTCCTAAAGCTGCTCGGCGAGAGTTTG 71
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 72 GACCGGGGCTCGCTCGGAGCGCTCCT -AAGCTGCTGCTCGGCGAGAGTTTG 130
OY 72 GAGGGGGGCTCGGAGAGTTTCCGTGGGGGCTCGGACGG -GACCTCGAGTTCG 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131 GAGGGGGGCTCGGAGAGTTTCCGTGGGGGCTCGGACGG -GACCTCGAGTTCG 190
OY 130 GCTTAAAGCTGCTCGGAGAGTTTCCGTGGGGGCTCGGAGCGGCTG -ACTGGCGACCTCGGAGC 188
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 191 GCTTAAAGCTGCTCGGAGAGTTTCCGTGGGGGCTCGGAGCGGCTG -ACTGGCGACCTCGGAGC 250
OY 189 CTGGCGGAGCGGAGCGCTCGGAGAGCGGAGACGCCCTTTTCCAGCAGCTG 248
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 251 GTGCGAGCGGCGAGCGCTCGGAGAGTCCGAGACGCCCTGCTCCCTCGACCGAGTG 310
OY 249 CTAGGCTTGGAGCCACAGAAAAAAGTAGAGTCCGGCTGCTTTCAGAGCTTG 308
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 311 -TAGTTTCGGAGACCTGCGGGGCGAAAGTAGAGTCCAGCGGCTTCAGCGCTTG 369
OY 309 ACGGCGGCGCGCTGGGAGAGAGTAGAGCGACCGCTTACCTAAAGTAGAAGCTG 368
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 370 ACGGCGGCGCGCTGGGAGAGAGTAGAGCGACCGCTTACCTAAAGTAGAAGCTG 429
OY 369 GGGTTGGTAGCCCTACTTTTGGGGGCTCGTGGAGAGCTGCTGGGCTCGAAGAGCA 428
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 430 GGGTTGGTAGCCCTACTTTTGGGGGCTCGTGGAGAGCTGCTGGGCTCGAAGAGCA 489
OY 429 AAGTCTACACTGTGAGAGTTGCGAGGGCTCTGCTGAGTAGAATTAGAGTGGAA 488
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 490 GAGTCTCCACTGTGAGAGTAGAGGGCTCTGCTGAGTAGAATTAGAGTGGAA 549
OY 489 CGTGAAGCCCAAGAACCTTACAGATGGAGTCTTCCAGATCAATCCAGATGGCA 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 550 GGTGAGCCCAAGAACCTTACAGATGGAGTCTTCCAGATCAATCCAGATGGCA 609
OY 549 GTGAGTGTGGAGAGTCTTATGCCGCTCAGAGGCCCACTCAGAGAGTTGTTGAG 608
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 610 GTGAGTGTGGAGAGTCTTATGCCGCTCAGAGGCCCACTCAGAGAGTTGTTGAG 669
OY 609 GGTGTGTGACCAATAGAGAGTACGGGGAGACAGATTGACCTTTCACCCAGAA 668
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 670 GATATGTACCGGATAGAGAGTATGGGGAGACAGATTGATCTTCCACCCATG 729
OY 669 CTACGTACGGGTGTGAGCGGGAATGAGATGCACTGAACTTACAGGCGATCCG 728
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 730 CTACGTACGTGTATGGCGCGGAGATGAGATCAGATGAACTGAGACCTCAGAA 789
OY 729 AATTGACTGAGATATGAGCGGACCCCTCAAGTTTGCCTGTGAGAGCATTTGAG 788
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 790 AATGCACTCAGATATTAGCGGCGACCTCAAGTTTGCCTGTGAGAGCATTTGAG 849
OY 789 CGAGGATGAGCTTATGCAATTTCTTTCAGAGAGGCTGACACGTTTAAAGCAA 848
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 850 CGAGGATGAGCTTATGCAATTTCTTTCAGAGAGGCTGACACGTTTAAAGCAA 909
OY 849 CACTAAGCGGAGATCTATGATGATGACATGCGGACAGATCTCAGAGATGATGAAT 908
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 910 CACTAAGCGGAGATCTATGATGATGACATGCGGACAGATCTCAGAGATGATGA 969
OY 909 CACTGAGGAGAGCGCTACCAACCAAGTAGATGAAACACCCCGAGAGGAGATGCA 968
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 970 CACTGAGGAGAGCGCAC - - - - -ACTGGCTTGTAGATGATCACCAGAGAG 1024
OY 969 GCATTGCGCTTT -TATATTACGTTTATGGAATGAACTGAAAAAA 1013
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1025 GCATGCGCTTTTATATTATATGTTTACTGAAATTAAGTGAAGAAAA 1071

```

RESULT 5

```

AX464040
LOCUS AX464040 1210 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 173 from Patent WO0140466.
ACCESSION AX464040
VERSION AX464040.1 GI:21899037
KEYWORDS
SOURCE
ORGANISM human.
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Baker, K.P., Beresini, M., DeForge, L., Desnoyers, L., Filvaroff, E.,
  Gao, M.Q., Gerritsen, M.E., Goddard, A., Godowski, P.J., Gurney, A.L.,
  Sherwood, S., Smith, V., Stewart, T.A., Tumas, D., Watanabe, C.K.,
  Wood, W.L. and Zhang, Z.
  Secreted and transmembrane polypeptides and nucleic acids encoding
  same
  JOURNAL Patent: WO 0140466-A 173 07-JUN-2001;
  Genentech Inc. (US)
FEATURES
    source
        1..1210
        Location/Qualifiers
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 261 a 316 c 379 g 254 t
ORIGIN
Query Match 64.6%; Score 691; DB 6; Length 1210;
Best Local Similarity 84.4%; Pred. No. 2e-176;
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;
OY 12 GAGGGGGGCTCGCTCGGAGCGCTCCTAAAGCTGCTCGGCGAGAGTTTG 71
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196 GACCGGGGCTCGCTCGGAGCGCTCCT -AAGCTGCTGCTCGGCGAGAGTTTG 254
OY 72 GAGGGCGGCGCTTAAAGGTAGTTTGGTGGGGGCTCGGACGG -GACCTCGAGTTCG 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 255 GAGGGCGGCTTGGGCTGCTGCTGATTTGGGCTCGACCGCACCGAGTTCG 314
OY 130 GCTTAAAGCTGCTCGGAGAGTGGGAGAGTGGGCGGAGGCTG -ACTGGCGACCTGAGC 188
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 315 GCTTAAAGCTGCTCGGAGAGTGGGAGAGTGGGCTGCGGCGGAGACCCGCTG 374
OY 189 CTGGCGGAGCGGAGCGCTCTGAGAGAGCGGAGACGCCCTTTTCCAGCAGCTG 248
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 375 GTGGCAGGCGGCGAAGCGCTGAGAAATCCCGAGACGCCCTGCTCCAGACCGAGTG 434
OY 249 CTAGGCTTGGAGCCACAGAAAAAAGTAGAGTCCGCTGCTTTCAGAGAGCTGG 308
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435 -TAGTTTCGGAGACCTGCGGCGCAAGTAGAGTCCAGCGGCTTCAGCGGCTTG 493
OY 309 ACGGCGGCGCGCTGGGAGCAGAGTGAAGCAAGCTGTTTACCTAAAGTAGAAGCTG 368
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 494 ACGGCGGCGCGCTGGGAGCAGAGTGAAGCAAGCTTACCTTAAAGTAGAAGCTG 553
OY 369 GGGTGGCTAGCCCTACTTTTGGGGGCTCGTGGAGTCCCTGCGGCTCGAAGAGCA 428
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 554 GGGTGGCTAGCCCTACTTTTGGGGGCTCGTGGAGTCCCTGCGGCTCGAAGAGCA 613
OY 429 AGATCTACACTGTGAGAGTTCGAGGGCTGCTGAGTAGAATTAGATGGGAATTTCC 488
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 614 GATATGTACCGGATAGAGAGTATGGGAGTAGAATGGAATTTGCCCA 673
OY 489 CGTGAAGCCCAAGAACCTTACAGATGGAGTCTTCCAGATCAATCCAGATGGAGCA 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 674 GGTGAGCCCAAGAACCTTACAGATGGAGTCTTCCAGATCAATCCAGATGGAGCA 733
OY 549 GTGAGTGTGGAGAGTATGATGCCGCTCAGAGGCCCACTCAGAGAGTTGCTTGA 608
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 734 GTGAGTGTGGAGAGTATGATGCCGCTCAGAGGCCCACTCAGAGAGTCTGGA 793
OY 609 GGTGTGTGACCGAATGAGAGTGAAGGAGGAGAAAGATTGACCTTACCCAGCGCAAG 668
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 794 GATATGTACCGGATAGAGAGTATGGGAGAACAGATTGATCTTCCACCCATGCAAG 853

```





CDS /clone\_lib="pkal-meta-1"  
145..693

/codon\_start=1  
/product="type II membrane protein"  
/protein\_id="BAA76498.1"  
/db\_xref="GI:4586840"

/translation="MKGMWLALLIGALITAMARRSODLHGCACRAVLDELMEIAQ  
VPPKTIQMSFRINPDQSIVPEVPAHSEHLEICDPAKREGEQIDPSTHR  
KNYRVNNGESSELDLQIRIDISDLSIGLKFACSEIYEVEEDLIEFFSEADNVK  
DKLSKRDLCDLHAHSHDEL"

BASE COUNT 210 a 194 c 241 g 169 t  
ORIGIN

## Query Match

Best Local Similarity 87.3%; Score 606.6; DB 9; Length 814;  
Matches 701; Conservative 0; Mismatches 94; Indels 8; Gaps 3;

OY 213 AGAAGCCGGGAGAGCCCTTTTCCACGAGCGTGTAGGGTTGGAGCCACAGAAAAC 272  
DB 1 AGAATCCCGAGAGCCCTCTCCCGACCGAGTG-1AGTTCCGGAGCCACTGGGGCC 59  
OY 273 AAGTGAAGTCCGCGCTGTTCCAGAGCTGGCCAGCGCGCGCGTGGAGACAG 332  
DB 60 AAGTGAAGTCCAGCGGCTTCCAGCGCTGGCCAGCGCGCGCGCGTGGAGACAG 119  
OY 333 GTGAGAGCAGCCTGTACACTAAGATGAAGGCTGGGTTGGCTAGCCCTACTTTGGG 392  
DB 120 GTGAGAGCAGCCTGTACACTAAGATGAAGGCTGGGTTGGCTAGCCCTACTTTGGG 179  
OY 393 GGTCTGCTGGGAGTCCCTGGGCTGAGAGAGCCAGATCTACCTGTGAGCTTGCAG 452  
DB 180 GGCCTGCTGGGAGTCCCTGGGCTGAGAGAGCCAGATCTCCACTGTGAGCTTGCAG 239  
OY 453 GGCCTGCTGGTGAATTAAGATGGGAATTCGCCGCTGAGACCCCAAGAACCATTC 512  
DB 240 GGCCTGCTGGTGAATTAAGATGGGAATTCGCCGCTGAGACCCCAAGAACCATTC 299  
OY 513 GATGGGATCTCTCCGAATCAATCCAGATGGCAGCAGTCACTGTGGAGTAACTTATGC 572  
DB 300 GATGGGATCTCTCCGAATCAATCCAGATGGCAGCAGTCACTGTGGAGTAACTTATGC 359  
OY 573 CCGGTGAGAGCCCACTGACAGAGTGGTGTGAGAGCTGTGAGACCGAATGAAGAGTA 632  
DB 360 CCGGTGAGAGCCCACTGACAGAGTGGTGTGAGAGCTGTGAGACCGAATGAAGAGTA 419  
OY 633 CGGGGAACAGATTGACCTTCTACCCACCGCAAGAACTAGAGCTCGTGAACCGGA 692  
DB 420 TGGGGAACAGATTGATCTTCCACCCATCCGAGAACTAGAGTGTAGTGGCCGGA 479  
OY 693 TGGGAATCCAGTGAATTAAGAGGCAATCCGAATTCAGATATACAGCGGCAC 752  
DB 480 TGGGAATCCAGTGAATTAAGAGGCAATCCGAATTCAGATATACAGCGGCAC 539  
OY 753 CTTCAAGTTCGCTGTGAGAGCAATGGGAAGATACAGAGATGAGCTTATCGAATTC 812  
DB 540 CTTCAAGTTCGCTGTGAGAGCAATGGGAAGATACAGAGATGAGCTTATCGAATTC 812  
OY 813 CTCGAGAGAGCTGACACGTTAAAGCAAACTTTCAGTAAGCGGACGATATATGTA 872  
DB 600 TTTCCGAGAGCTGACAAATGTTAAAGCAAACTTTCAGTAAGCGGACGATATGTA 659  
OY 873 CCAATGCCCTGACAGATTCACAGATGATGATGATGATGATGATGATGATGATGAT 932  
DB 660 CCAATGCCCTGACAGATTCACAGATGATGATGATGATGATGATGATGATGATGAT 714  
OY 933 AACGTATGAACACCCCGAGAGGAGATGAGCAATTCGCTT--TATATTACGTT 990  
DB 715 GGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774  
OY 991 TTTATGGAATTAAGTGAAGAAA 1013  
DB 775 TTTACTGAATTAAGTGAAGAAA 797

## RESULT 8

AC090489

## LOCUS

AC090489

## DEFINITION

AC090489

## ACCESSION

AC090489

## KEYWORDS

AC090489

## SOURCE

AC090489

## ORGANISM

AC090489

## TITLE

AC090489

## JOURNAL

AC090489

## REFERENCE

AC090489

## AUTHORS

AC090489

## JOURNAL

AC090489

## REFERENCE

AC090489

## AUTHORS

AC090489

## JOURNAL

AC090489

## COMMENT

AC090489

## FEATURES

AC090489

## source

AC090489

## misc-feature

AC090489

## BASE COUNT

AC090489

## ORIGIN

AC090489

## Query Match

AC090489

## Best Local Similarity

AC090489

## Matches 321; Conservative

AC090489

## 0; Mismatches 0; Indels 0; Gaps 0;

## OY 16 GGGGCTCCGCTGCTGCGAGGCGCTCTTAAGCTGCTGCTGCGGAGAGATTGGAGG 75

## DB 72184 GGGGCTCCGCTGCTGCGAGGCGCTCTTAAGCTGCTGCTGCGGAGAGATTGGAGG 72243

## OY 76 GCGGCGCTTAAGGCTGCTGCGAGGCGCTGCGAGGCGCTGCGAGGCGCTGCGAGG 135

## DB 72244 GCGGCGCTTAAGGCTGCTGCGAGGCGCTGCGAGGCGCTGCGAGGCGCTGCGAGG 72303

## OY 136 GTGCTATTTAAGTGGGAGGCTGCGAGGCGCTGCGAGGCGCTGCGAGGCGCTGCGAGG 195

## DB 72304 GTGCTATTTAAGTGGGAGGCTGCGAGGCGCTGCGAGGCGCTGCGAGGCGCTGCGAGG 72363

## OY 196 GAGCGGAGCGCTGCGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 255

```

|||||
DB 72364 GAGCCGAGACGCTCTGAGAGACCGGAGACGCCGCTTTTCCACAGCTGTAGGGT 72423
OY 256 TGGGACCCACAGAAAACAAAGTAGAGATCCGGCTCTTTCCAGAGCCTGGCCACGGCGG 315
DB 72424 TGGGACCCACAGAAAACAAAGTAGAGATCCGGCTCTTTCCAGAGCCTGGCCACGGCGG 72483
OY 316 CGGCGCTGGGACGACGAGTGG 336
DB 72484 CGGCGCTGGGACGACGAGTGG 72504

RESULT 9
BC001027 824 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, transmembrane protein 4, clone MGC:1545
IMAGE:344788, mRNA, complete cds.
BC001027
ACCESSION BC001027.1 GI:12654402
VERSION MGC.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 824)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov/
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nrl.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantirip, S., Thomas, P.J.,
Tjongson, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IMAGE Plate: 5 Row: P Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6014631.

FEATURES
Source
1..824
/organism="Homo sapiens"
/db_xref="locusid:10330"
/db_xref="taxon:9606"
/clone="MGC:1545 IMAGE:344788"
/tissue_type="Eye, retinoblastoma"
/clone_id="NH_MGC_16"
/lab_host="DH10B-R"
/Note="Vector: pOTB7"
250..504
/codon_start=1
/product="transmembrane protein 4"
/protein_id="AAH01027.1"
/db_xref="gi:12654403"
/translators="MKMGWLAALGALGTAAARSDLLCGACRALVDELEWELAQ
VDPKKTIONSGFRINPDGSSOSVETVTYPRNKVAHSGFC"

BASE COUNT 221 a 194 c 233 g 176 t

```

```

ORIGIN
Query Match 28.7%; Score 307; DB 9; Length 824;
Best Local Similarity 82.9%; Pred. No. 3.e-72;
Matches 374; Conservative 0; Mismatches 75; Indels 2; Gaps 2;

OY 123 GATCTCCGCTTAGGTGCTTAACTGAAGTGGGAGAGTGGCCAGCGGTC-ACTGGCCAC 181
DB 15 GAGTCCCGCTTAGGTGCTTAACTGAAGTGGGAGAGTGGCCAGCGGTC-ACTGGCCAC 74
OY 182 CCGTAACCTGGGCGGAGCGCGCTTGGAGAACCGCGGAGCGCGGCTTTTCCCGAG 241
DB 75 CCGGTGTGGCAGCGCGGAGCGCGCTTGGAGAACCGCGGAGCGCGGCTTTTCCCGAG 134
OY 242 CCAGCTGTAGGTGGGAGCCACAGAAAACAAAGTAGAGTCCGCTCTTTCCAGAC 301
DB 135 CCAGGTG-TAGTGTGGGAGCCACTGGGGCCAAAGTAGAGTCCAGGGTCTTCCAGCGC 193
OY 302 CTGGGCCACGCGCGCGCGCGCTGGGAGCAGAGTGGAGCCACCTGTACATTAAGATGA 361
DB 194 TTGGGCCACGCGCGCGCGCGCTGGGAGCAGAGTGGAGCCACCTGTACATTAAGATGA 253
OY 362 AAGGCTGGGGTGGCTTACCTTTTGGGGTCTGCTGGGAACCTGGCGCTCGAA 421
DB 254 AAGGCTGGGGTGGCTTACCTTTTGGGGTCTGCTGGGAACCTGGCGCTCGGA 313
OY 422 GGAGCCAGATCTACACTGTGGAGCTTGCAGGCGCTGTGTGATGAATAGAGTGGAAA 481
DB 314 GGAGCCAGATCTTCCACTGTGGAGCAGTGCAGGCGCTGTGTGATGAATAGAGTGGAAA 373
OY 482 TTGGCGCGTGGAGCCCAAGAACCATTCAGATGGATGCTTCCGAATCAATCCAGATG 541
DB 374 TTGGCGCGTGGAGCCCAAGAACCATTCAGATGGATGCTTCCGAATCAATCCAGATG 433
OY 542 GCAGCAGTCACTTGGGAGTACCTTATG 572
DB 434 GCAGCAGTCACTTGGGAGTACCTTATG 464

RESULT 10
AC109891 10166 bp DNA linear HTG 13-JUL-2002
LOCUS Rattus norvegicus clone CH230-31966, *** SEQUENCING IN PROGRESS
DEFINITION *** 47 unordered pieces.
AC109891
AC109891.3 GI:21738196
VERSION HTG; HTGS; PHASE1.
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 101666)
AUTHORS Munny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
Albrooks, S.L., Amaralunge, H.C., Are, J.R., Ayala, M., Banks, T.,
Barbarta, J., Benton, J., Blincoe, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowler, S., Bileva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carroll, T.F., Carter, M., Cavazos, R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
DeLaney, K.R., Davis, C., Davy-Carroll, L., Dederich, D.A.,
DeVella, M.L., Davis, C., Davy-Carroll, L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhardt, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, D., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorell, J., H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlik, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hughes, M., Holloway, C., Hollins, B.,
Honsi, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivel, S., Joudah, S.,
Karlisson, E., Kelly, S., Khan, U., King, L., Korvah, D., Kovar, C.,

```

Kralovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Louieged, H.,  
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M.,  
Miner, G., Miner, Z., Mitchell, J., Mohabat, K., Morgan, K., Morris, S.,  
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokenko, S., Ogih, M., Okwunnu, G.,  
Oragunye, N., Oriedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rivera, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,  
Scherrer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,  
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Taber, T., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansley, J., Taylor, C., Taylor, T., Teitrod, B., Thomas, N., Thomas, S.,  
Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, O.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Nelson, D.,  
Weinstock, G. and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 101666)  
Morley, K.C.

Direct Submission  
Submitted (08-FEB-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 101666)  
Morley, K.C.

Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 12, 2002 this sequence version replaced gi:18860222.

----- Genome Center -----  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information -----  
Center project name: GOKY  
Center clone name: CH230-319P6  
----- Summary Statistics -----  
Sequencing vector: Plasmid  
Assembly: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap: version 0.990329  
Consensus quality: 60229 bases at least Q40  
Consensus quality: 63959 bases at least Q30  
Consensus quality: 66730 bases at least Q20

----- NOTE: Estimated insert size may differ from sequence length -----  
\* (see <http://www.hgsc.bcm.tmc.edu/docs/genbankdraft.html>).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 47 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1423: contig of 1423 bp in length  
\* 1424 1523: gap of unknown length  
\* 1524 2972: contig of 1449 bp in length  
\* 2973 3072: gap of unknown length  
\* 3073 4717: contig of 1645 bp in length  
\* 4718 4817: gap of unknown length  
\* 4818 6632: contig of 1815 bp in length  
\* 6633 6732: gap of unknown length  
\* 6733 8040: contig of 1308 bp in length  
\* 8041 8140: gap of unknown length  
\* 8141 10043: contig of 1903 bp in length  
\* 10044 10143: gap of unknown length

10144 11237: contig of 1094 bp in length  
\* 11238 11337: gap of unknown length  
\* 11338 12362: contig of 1025 bp in length  
\* 12363 12462: gap of unknown length  
\* 12463 14100: contig of 1538 bp in length  
\* 14101 14200: gap of unknown length  
\* 14201 15840: contig of 1640 bp in length  
\* 15841 15940: gap of unknown length  
\* 15941 17053: contig of 1113 bp in length  
\* 17054 17153: gap of unknown length  
\* 17154 18836: contig of 1683 bp in length  
\* 18837 18936: gap of unknown length  
\* 18937 20137: contig of 1201 bp in length  
\* 20138 20237: gap of unknown length  
\* 20238 21736: contig of 1499 bp in length  
\* 21737 21837: gap of unknown length  
\* 21837 23620: contig of 1783 bp in length  
\* 23620 23720: gap of unknown length  
\* 23720 25907: contig of 2188 bp in length  
\* 25907 26007: gap of unknown length  
\* 26008 27406: contig of 1399 bp in length  
\* 27407 27506: gap of unknown length  
\* 27507 28837: contig of 1331 bp in length  
\* 28838 28937: gap of unknown length  
\* 28938 30229: contig of 1292 bp in length  
\* 30230 30329: gap of unknown length  
\* 30330 32012: contig of 1683 bp in length  
\* 32013 32112: gap of unknown length  
\* 32113 33198: contig of 1086 bp in length  
\* 33199 33298: gap of unknown length  
\* 33299 35519: contig of 2221 bp in length  
\* 35520 35620: gap of unknown length  
\* 35620 37678: contig of 2038 bp in length  
\* 37678 37777: gap of unknown length  
\* 37777 38936: contig of 1159 bp in length  
\* 38937 39036: gap of unknown length  
\* 39037 40273: contig of 1237 bp in length  
\* 40274 40373: gap of unknown length  
\* 40374 41906: contig of 1533 bp in length  
\* 41907 42006: gap of unknown length  
\* 42007 44507: contig of 2500 bp in length  
\* 44507 44606: gap of unknown length  
\* 44607 46803: contig of 2197 bp in length  
\* 46804 46903: gap of unknown length  
\* 46904 49226: contig of 2323 bp in length  
\* 49227 49326: gap of unknown length  
\* 49327 51433: contig of 2107 bp in length  
\* 51434 51533: gap of unknown length  
\* 51534 53763: contig of 2230 bp in length  
\* 53764 53863: gap of unknown length  
\* 53864 55613: gap of unknown length  
\* 55614 56713: contig of 2750 bp in length  
\* 56714 58773: gap of unknown length  
\* 58774 58873: contig of 2060 bp in length  
\* 58874 61527: gap of unknown length  
\* 61528 61627: contig of 2654 bp in length  
\* 61628 63192: gap of unknown length  
\* 63193 63292: contig of 1565 bp in length  
\* 63293 65362: gap of unknown length  
\* 65363 65462: contig of 2070 bp in length  
\* 65463 67438: gap of unknown length  
\* 67439 67538: contig of 1976 bp in length  
\* 67539 70505: gap of unknown length  
\* 70506 70605: contig of 2967 bp in length  
\* 70606 74524: gap of unknown length  
\* 74525 74624: contig of 3919 bp in length  
\* 74625 76737: gap of unknown length  
\* 76738 76837: contig of 2113 bp in length  
\* 76838 80133: gap of unknown length  
\* 80134 80233: contig of 3296 bp in length  
\* 80234 83122: gap of unknown length  
\* 83123 83222: contig of 2889 bp in length  
\* 83223 86731: gap of unknown length  
\* 86732 86733: contig of 3509 bp in length

\* 86732 86831: gap of unknown length  
\* 86832 89929: contig of 3098 bp in length  
\* 89930 90029: gap of unknown length  
\* 90030 93107: contig of 3078 bp in length  
\* 93108 93207: gap of unknown length  
\* 93208 97419: contig of 4212 bp in length  
\* 97420 97519: gap of unknown length  
\* 97520 101666: contig of 4147 bp in length.  
Location/Qualifiers  
1. 101666  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-319B6"  
BASE COUNT 24650 a 22934 c 23569 g 23059 t 7454 others  
ORIGIN

Query Match 23.3% Score 248.6; DB 2; Length 101666;  
Best Local Similarity 90.7%; Pred. No. 5.5e-56;  
Matches 291; Conservative 0; Mismatches 19; Indels 11; Gaps 2;

QY 16 GGGGCTCTGCTGCTCGAGAGCGCTCTTAAGCTGCTGCTGCGCGAGAGTTGGAGG 75  
|||||  
Db 72973 GGGGCTCTGCTGCTCGAGAGCGCTCTTAAGCTGCTGCTGCGCGAGAGTTGGAGG 73032  
|||||

QY 76 GGGGGCTTAGGCTAGTTGCTGGGGGCTCGCAGCGAGACCTCAATCTCCGCTTAG 135  
|||||  
Db 73033 GGGGCTTAGGCTAGTTGCTGGGGGCTCGCAGCGAGACCTCAATCTCCGCTTAG 73090  
|||||

QY 136 GTGCTTAGTTAAGTCGGGAGAGCTGGGCGGCTGCACTGCGCAGCCTGAAGCTGGCGG 195  
|||||  
Db 73091 GTTCTGTAGTACGGGGGCTGCTGGGCGGCTGCACTGCGCAGCCTGAAGCTGGCGG 73150  
|||||

QY 136 GAGCGGAGAGCTCTGGAGAGCGCGGACAGCCCGTTTCCAGACCAAGCTGTAAGGT 255  
|||||  
Db 73151 GCGCGGAGAGCTCTGGAGAGCGCGGACAGCCCGTTTCCAGACCAAGCTGTAAGGT 73210  
|||||

QY 256 TGGGACCCACAGAAACAAAGTAGAGTCCGGCTTCCAGAGCCGCGGCGAGCGGCG 315  
|||||  
Db 73211 TGGGAGC-----CAAAGTAGAGTCCGGCTTCCAGAGCCGCGGCGAGCGG 73261  
|||||

QY 316 CGGCGCTGGGAGCAGAGGTGG 336  
|||||  
Db 73262 CGGCGCTGGGAGCAGAGGTGG 73282  
|||||

RESULT 11  
AF186113 649 bp mRNA linear PRI 13-JAN-2000  
LOCUS Homo sapiens putative secreted protein ZS1G9 (ZS1G9) mRNA, complete cds.  
DEFINITION  
ACCESSION AF186113  
VERSION AF186113.1 GI:6014631  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 649)  
AUTHORS Sheppard,P., Jelinek,L., Whitmore,T., Blumberg,H., Lehner,J. and O'Hara,P.  
TITLE Homo sapiens putative secreted protein  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 649)  
AUTHORS Sheppard,P., Jelinek,L., Whitmore,T., Blumberg,H., Lehner,J. and O'Hara,P.  
TITLE Direct Submission  
JOURNAL Submitted (14-SEP-1999) Biomolecular Informatics, ZymoGenetics, Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA  
FEATURES  
source  
1. 649  
/organism="Homo sapiens"  
/db\_xref="dbEST:R15306"  
/db\_xref="taxon:9606"

/chromosome="12"  
/map="12q15"  
/clone="IMAGE:29675"  
1. 649  
/gene="ZS1G9"  
104. 358  
/gene="ZS1G9"  
/codon\_start=1  
/product="putative secreted protein ZS1G9"  
/protein\_id="AA01431.1"  
/db\_xref="GI:6014632"  
/translation="MKGGWALLIGALITGMARRSDLLGCAALVDELLEWELIAQ  
VDPKKTIDMSFRINPDSSQSVETVTVPNKAHSGEG"  
BASE COUNT 175 a 150 c 171 g 149 t 4 others  
ORIGIN

Query Match 23.1% Score 246.8; DB 9; Length 649;  
Best Local Similarity 88.4%; Pred. No. 7e-56;  
Matches 266; Conservative 1; Mismatches 34; Indels 0; Gaps 0;

QY 272 CAAAGTGAAGTCCGGCTGCTTCCAGAGCCTGGGCGGCGGCGGCGGCGGAGACAGA 331  
|||||  
Db 18 CAAAGTGAAGTCCAGCGGCTGCTTCCAGAGCCTGGGCGGCGGCGGCGGCGGAGACAGA 77  
|||||

QY 332 GGTGAGCGACCTGTTACACTAAAGAGTAAAGGCTGGGCTGGCTAGCCCTACTTTGG 391  
|||||  
Db 78 GGTGAGCGAACCCCGTTACCTAAATGAATGAAGGCTGGGCTGGGCTGGCTAGCCCTACTTTGG 137  
|||||

QY 392 GGGCTCTGCTGGGAACTGGCTGGGCTGGAAGGAGCCAAATGATCTACACTGTGAGCTTGA 451  
|||||  
Db 138 GGGCTCTGCTGGGAACTGGCTGGGCTGGAAGGAGCCAAATGATCTACACTGTGAGCTTGA 197  
|||||

QY 452 GGGCTCTGCTGGGAACTGGGCTGGGCTGGAAGGAGCCAAATGATCTACACTGTGAGCTTGA 511  
|||||  
Db 198 GGGCTCTGCTGGGAACTGGGCTGGGCTGGAAGGAGCCAAATGATCTACACTGTGAGCTTGA 257  
|||||

QY 512 AGATGGATCTCTCCGATCAATCCAGATGGCAGCAGTCACTTTGGAGGTAAGTATG 571  
|||||  
Db 258 AGATGGATCTCTCCGATCAATCCAGATGGCAGCAGTCACTTTGGAGGTAAGTATG 317  
|||||

QY 572 C 572  
|  
Db 318 C 318  
|

RESULT 12  
AX440375 564 bp DNA linear PAT 28-JUN-2002  
LOCUS AX440375  
DEFINITION Sequence 228 from Patent WO0190154.  
ACCESSION AX440375  
VERSION AX440375.1 GI:21665185  
KEYWORDS  
SOURCE  
ORGANISM human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Xu,J., Mitcham,V.L., Harlocker,S.L., Dillon,D.C., Secrist,H., Lodes,M.J., Algate,P.A., Fling,S.P., Mannion,J., Benson,D.R. and Carter,D.  
TITLE Compositions and methods for the therapy and diagnosis of ovarian cancer  
JOURNAL Patent: WO 0190154-A 228 29-NOV-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source  
1. 564  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 147 a 142 c 152 g 118 t 5 others  
ORIGIN

Query Match 20.7% Score 221; DB 6; Length 564;  
Best Local Similarity 83.2%; Pred. No. 6.9e-49;

	Matches	298;	Conservative	0;	Mismatches	50;	Indels	10;	Gaps
QY	211	GGAGAACCCGGGACACAGCCCCGTTTTTTTCCAGCCAGCTGCTAGAGGTTGGAGCCACAGAAA							270
Db	216	GGGAATCCCGGACACCCCTGCTCCCTCGGACAGGTG-TAGTTTTGGAGGCCACCTGGGG							274
QY	271	ACAAATGAGAGTCCGGGTGCTTTCAGAGCCTGGGCCACGGCGGGCCGTGGGAGCAG							330
Db	275	CCAAATGTAGATCCAGCCGCTTTCAGGCGCTTGGGCCACGGCGGGCCCTGGGAGCAG							334
QY	331	AGGTGAGCGGACCCCTGTATTACACTAAAGATGAAGAAGCTGGAGTTGGCTAGCCCTACTTTTG							390
Db	335	AGGTGAGGAGACCCCATTTACGCTAAAGATGAAGAAGCTGGGTTGGCTGACCTCGTTCTTG							394
QY	391	GGGTCCTGCTGGGGAACGCGCTGGGCTCGAAGGACCCAAAGATCTACACTGTGAGACTTGC							450
Db	395	GGGCGCCCTGCTGGGAACCGCCTGGGCTGGAGGACCCAGAGATCTCACCTGTG-----							447
QY	451	AGGGCTCTGGTGATGAATTAGATGGGGAATTGGCCCGCGTGGAGCCCAAGAAGACATT							510
Db	448	AGGGCTCTGTGGATGAAGTGAATGGAATTG-CCANGTGGACCCCAAGAAGACATT							506
QY	511	CAGATGGGATCTTCCGGAATC-AATCCAGATGGGACCGCAGTTCAGTTGGAGATACCG							567
Db	507	CAGAAAGGATCTTCCGATCAATTCAGATGGGACCAACGATCAATGGTGGANGGCCCT							564

Sutton, A., Svatik, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansley, C., Taylor, C., Taylor, T., Telifod, B., Thomas, N., Thomas, S.,  
 Uman, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
 Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 155023)  
 Worley, K.C.  
 Direct Submission  
 Submitted (11-MAR-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 155023)  
 Worley, K.C.  
 Direct Submission  
 Submitted (09-MAY-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On May 3, 2002 this sequence version replaced v1.2022551.

RESULT	13
AC025574/c	
LOCUS	
DEFINITION	15023 bp DNA linear HTG_09-MAY-2002
AC025574	
Homo sapiens chromosome 12 clone RP11-348M3, WORKING DRAFT SEQUENCE, 12 unordered pieces.	
AC025574	
AC025574.13	GI:20428723
HTG: HTGS_PHASE1; HTGS_DRAFT.	
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens

REFERENCE  
1 (bases 1 to 155023)  
Muzny, D.M., Adams, C., Adto-Gdula, B. et al. 1996. The human genome: initial sequencing and assembly. *Science* 273: 1503-1512.

Barbára, Benton, B. J., Bontlinger, A. C., Aye, R., Ayele, M., Banks, T., Bonick, J., Boyle, S., Bilewa, K., Blumke, K., Blankenburg, K., Bonatti, D., Bunyah, C., Burch, P., Burnett, C., Burrell, K. L., Byrd, N. C., Carroon, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, M. C., Cox, C., Coyle, M. D., Dathorne, S. R., David, R., Davila, M. L., Davis, C., Day-Carroll, L., Dederich, D. H., Delnawate, K. R., Delgado, O., Denn, A. L., Ding, Y., Dinh, H. H., Eamhardt, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Galati, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gottrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, J., Homsl, F., Howard, S., Hubber, J., Hulys, S., Hume, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jollivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Kovach, J., Kovar, C., Kralovic, J., Kuresh, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L., Li, J., Li, Z., Lichner, O., Iken, C., Liu, J., Liu, W., Lounsbeged, H., Lozadó, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Matthez, E., Massey, E., Mawhin, E., McLeod, M. P., Meddow, M., Mel, G., Metzger, S., Minter, G., Mner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, M., Moser, M., Neal, D., Newton, E., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nicholson, S., Nwokwenkwo, S., Ogulu, M., Okunolu, G., Oragunye, N., Oyedero, R., Pace, A., Payton, B., Peery, J., Perez, J., Peters, L., Pickens, R., Primus, E., Paul, L., Qules, M., Ren, Y., Rives, M., Rojas, A., Rojuckan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,

```

-----Vastly coverage: 4x in Q20 bases; sum-of-contigs estimation
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 12 contigs. The true order of the pieces
  is not known and their order in this sequence record is
  arbitrary. Gaps between the contigs are represented as
  runs of N, but the exact sizes of the gaps are unknown.
  This record will be updated with the finished sequence
  as soon as it is available and the accession number will
  be preserved.

1
3033: contig of 3033 bp in length
3034 3133: gap of unknown length
3134 5972: contig of 2838 bp in length
5972 6072: gap of unknown length
6072 10269: contig of 4198 bp in length
10270 10369: gap of unknown length
10370 13009: contig of 2640 bp in length
13010 13109: gap of unknown length
13110 17573: contig of 4464 bp in length
17574 17673: gap of unknown length
17674 24650: contig of 6977 bp in length
24651 24750: gap of unknown length
24751 31778: contig of 7028 bp in length
31779 31878: gap of unknown length
31879 42263: contig of 10385 bp in length
42264 43263: gap of unknown length
43264 51104: contig of 8741 bp in length
51105 51204: gap of unknown length
51205 68306: contig of 17102 bp in length
68307 68406: gap of unknown length
68407 109481: contig of 41075 bp in length
109482 109581: gap of unknown length

```



## ORIGIN

Query Match 16.6%; Score 177.4; DB 2; Length 182350;  
 Best Local Similarity 90.0%; Pred. No. 1.3e-36;  
 Matches 190; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

OY 555 TGTGAGGTACTTATGCGCCCTCAGAGCCCTCAGACAGTGTGTTGAGGAGTGTG 614  
 DB 172034 TTTCTAGGTGCTTATGCGCCCTCAGAGCCCTCAGACAGTGTGTTGAGGAGTGTG 614

OY 615 TGACCGAATGAAGAGTACGCGCAACAGATGACCTTACACCCAGCAAGTACGT 674  
 DB 171974 TGACCGGATGAAGAGTATGGGCAACAGATGATCTTCCACCATGCAAGTACGT 171915

OY 675 ACGGCTGCTGAGCCGCAATGAGATCCAGTACCTTACAGGCAATCCGCAATGCA 734  
 DB 171914 ACGGTGATGAGCGCGCAATGAGATCCAGTACCTTACAGGCAATCCGCAATGCA 171855

OY 735 CTCGATATCAGCGCACCTCAAGTTGCG 765  
 DB 171854 CTCGATATCAGCGCACCTCAAGTTGCG 171824

RESULT 15  
 AC023500 184762 bp DNA linear HTG 24-AUG-2002  
 LOCUS Homo sapiens chromosome 12 clone RP11-183H16, WORKING DRAFT  
 DEFINITION AC023500  
 ACCESSION AC023500.32 GI:22095070  
 VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 KEYWORDS Human.  
 SOURCE Human sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 184762)  
 Muzny D.M., Adams C., Adlo-Oduola B., Ali-Osman F.R., Allen C.,  
 Albrooke S.L., Amaralunge H.C., Are J.R., Ayele M., Banks T.,  
 Barakat J., Benton J., Bimage K., Blankenburg K., Bonin D.,  
 Boulikas J., Bowe S., Brileva M., Brown E., Brown M., Bryant N.P.,  
 Burch P., Burkett C., Burrell K.L., Byrd N.C.,  
 Carron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D.,  
 Chen G., Chen R., Chen Z., Chowdhury I., Christopoulos C.,  
 Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,  
 Davila M.L., Davis C., Davy-Carroll L., Dederich D.A.,  
 Delaney K.R., Delgado O., Denn A.L., Ding Y., Dinh H.H.,  
 Douthwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J.,  
 Eathart C., Edgar D., Edwards C.C., Elhaj C., Escotto M.,  
 Falls T., Ferraguto D., Flagg N., Ford J., Foster P., Frantz P.,  
 Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R.,  
 Correll J.H., Guevara W., Gunaratne P., Hale S., Hamilton K.,  
 Harris C., Harris K., Hart M., Havlak P., Hawes A., Hernandez J.,  
 Hernandez O., Hodgson A., Hogue M., Hollaway C., Hollins B.,  
 Homel F., Howard M., Huber J., Huiyk S., Hume J., Jackson L.E.,  
 Jacobson B., Jia Y., Johnson R., Jolivet S., Joudah S.,  
 Karlsson E., Kelly S., Khan U., King L., Korvah J., Kovar C.,  
 Kratovic J., Kureshi A., Landry N., Leal B., Lewis L.C., Lewis L.,  
 Li J., Li Z., Lichtarge O., Liu C., Liu Y., Liu W., Louised H.,  
 Lozano R.J., Lu X., Lucier A., Lucier R., Luna R., Ma J.,  
 Maheshwari M., Mapua P., Martin R., Martindale A., Martinez E.,  
 Massey E., Mauney E., McLeod M.P., Meador M., Mei G., Metzger M.,  
 Miner G., Miner D., Mitchell T., Mohabbat K., Morgan M., Morris S.,  
 Moser M., Nelson J., Newton E., Nwokwenkwo S., Oguh M., Okunolu G.,  
 Nguyen N., Nickerson E., Nwankwo S., Oguh M., Nguyen A., Nguyen N.,  
 Oragunye N., Oviedo R., Pace A., Payton B., Peery J., Perez L.,  
 Peters L., Pickens R., Pimus E., Pu L.L., Qulles M., Ren Y.,  
 Rivers M., Rojas A., Rojibokan I., Rolfe M., Ruiz S., Savary G.,  
 Scherer S., Scott G., Shen H., Shoostlari N., Sisson I.,  
 Sodergren E., Sonalke T., Sparks A., Stanley H., Stone H.,  
 Sutton A., Swalek A., Tabor P., Tamerisa A., Tamerisa K., Tang H.,  
 Tansey J., Taylor C., Taylor T., Telford B., Thomas R., Thomas S.,  
 Usmani K., Vasquez L., Vera V., Villalón D., Vinson R., Wang Q.,  
 Wang S., Ward-Moore S., Warren R., Washington C., Watlington S.,

## TITLE

JOURNAL  
 TITLE  
 AUTHORS  
 JOURNAL

REFERENCE  
 TITLE  
 AUTHORS  
 JOURNAL

## COMMENT

Williams G., Williamson A., Wlezyk R., Wooden S., Worley K.,  
 Wu C., Wu Y., Wu Y.F., Zhou J., Zorrilla S., Nelson D.,  
 Weinstein G., and Gibbs R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 184762)  
 Worley K.C.  
 Direct Submission  
 Submitted (15-FEB-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 184762)  
 Worley K.C.  
 Direct Submission  
 Submitted (24-AUG-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Aug 4, 2002 this sequence version replaced gi:22094227.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Drafting Center Code: BCM  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: HARC  
 Center clone name: RP11-183H16  
 ----- Summary Statistics  
 Sequencing vector: Plasmid  
 Sequencing vector: M13  
 Chemistry: Dye-Primer Body: 2% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 183845 bases at least Q40  
 Consensus quality: 184849 bases at least Q40  
 Estimated insert size: 185447 bases at least Q20  
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_drift\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html)).  
 NOTE: This is a "working draft" sequence. It currently  
 consists of 9 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1 2040: contig of 2040 bp in length  
 2041 2140: gap of unknown length  
 2141 4170: contig of 2030 bp in length  
 4171 4270: gap of unknown length  
 4271 8294: contig of 4024 bp in length  
 8295 8394: gap of unknown length  
 8395 14196: contig of 5802 bp in length  
 14197 14296: gap of unknown length  
 14297 33076: contig of 18780 bp in length  
 33077 33177: gap of unknown length  
 33178 52395: contig of 19119 bp in length  
 52396 52396: gap of unknown length  
 52397 84185: contig of 31790 bp in length  
 84186 84285: gap of unknown length  
 84286 120023: contig of 35738 bp in length  
 120024 120123: gap of unknown length  
 120124 184762: contig of 64639 bp in length.  
 Location/Qualifiers  
 1. 184762  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /clone="RP11-183H16"

## BASE COUNT

47425 a 44751 c 45101 g 46673 t 812 others





GenCore version 5.1.4\_p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 17:35:17 ; Search time 253.709 Seconds  
(without alignments)  
9488.757 Million cell updates/sec

Title: US-10-082-502-18

Perfect score: 1069  
Sequence: 1 gaattcgacagaggggggtc.....aaaaaatttcgcgcgcgc 1069

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

1:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1069	100.0	1069	20	AAK06970
2	1069	100.0	1069	21	AAZ08294
3	691	64.6	1210	22	AAK21330
4	676.4	63.3	1180	20	AAK22112
5	606.6	56.7	814	21	AAZ38327
6	606.6	56.7	814	24	ABK52765
7	606.6	56.7	814	24	ABK41995
8	606.6	56.7	814	24	ABK09772
9	599.6	56.1	806	20	AAK06969

10	599	56.0	806	21	AAZ08293	Human Zs199 gene e
11	564.6	52.8	832	20	AAK97837	Human secreted pro
12	555.2	51.9	714	20	AAK97884	Human secreted pro
13	513.4	48.0	592	21	AAK43525	Mouse secreted exp
14	503.2	47.1	657	24	ABK36007	CDNA sequence #398
15	472.4	44.2	546	21	AAZ38326	Human Transmembran
16	345.4	32.3	1085	20	AAK22130	WO9901020 Seq ID 3
17	281.2	26.3	792	20	AAK97836	Human secreted pro
18	256.2	24.0	718	20	AAZ10655	CDNA encoding a hu
19	253	23.7	649	20	AAK97883	Human secretory pe
20	246.8	23.1	649	20	AAK06968	Human secretory pe
21	246.6	23.1	649	21	AAZ08284	Human Zs199 gene.
22	243.4	22.8	558	22	AAZ18023	Human breast cance
23	238.4	22.3	415	20	AAK06971	Zs199 expressed se
24	237.8	22.2	415	21	AAZ08285	Expressed Sequence
25	233.2	21.8	595	22	AAZ07882	Human breast cance
26	231	20.7	564	24	ABK09691	Human ovarian tumo
27	210.8	19.7	484	22	AAZ25628	Human breast cance
28	210.8	19.7	1160	22	AAZ12689	Human breast cance
29	160.2	15.0	352	22	AAK09488	Human secreted pro
30	157.8	14.8	265	20	AAK39549	Human pancreatic c
31	155.6	14.6	2532	22	AAK07341	Human genomic DNA
32	155.6	14.6	2532	22	AAK32768	Human pancreatic c
33	155.6	14.6	2724	22	ABK07340	Human pancreatic c
34	155.6	14.6	2724	22	AAK32767	Human genomic DNA
35	153.6	14.4	495	22	AAZ180994	Human polynucleoti
36	84	7.9	215	21	AAC24962	Human secreted pro
37	66.8	6.2	147	24	ABK94837	Gene #1335 used to
38	66.8	6.2	147	24	ABK65469	Lung cancer relate
39	65	6.1	65	24	ABK55235	Mouse spliced tran
40	51.6	4.8	126	24	ABK94836	Gene #1334 used to
41	51.6	4.8	126	24	ABK65468	Lung cancer relate
42	51	4.8	60	24	ABK39205	Human spliced tran
43	49	4.6	467	23	ABK48839	Human prostate exp
44	48	4.5	475	23	ABK33270	Human prostate exp
45	48	4.5	475	23	ABK42193	Human prostate exp

## ALIGNMENTS

RESULT 1  
ID AAK06970 standard; cDNA; 1069 BP.  
AC AAK06970;  
XX  
XX 10-MAY-1999 (first entry)  
DT  
XX  
DE Mouse secretory peptide-9 (Zs199) orthologue cDNA.  
DE  
XX  
XX Secretory peptide-9; Zs199; orthologue; mouse; tumour marker;  
KW cancer; therapy; diagnosis; growth enhancer; ss.  
XX  
XX  
OS Mus sp.  
XX  
XX  
FH Key  
FT CDS  
FT  
FT sig\_peptide  
FT  
FT mat\_peptide  
FT  
FT  
PN WO9901554-A1.  
XX  
XX 14-JAN-1999.  
XX  
XX  
XX 02-JUL-1998; 98WO-US13859.  
XX  
XX 17-JUN-1998; 98US-0099005.  
XX 03-JUL-1997; 97US-0051704.  
XX 03-JUL-1997; 97US-0888088.

Query Match	100.0%	Score 1069	DB 20	Length 1069
Best Local Similarity	100.0%	Pred. No. 8.9e-268		
Matches 1069	Conservative 0	Mismatches 0	Indels 0	Gaps 0
1	GAATTCGGCAGAGGGGGTCTCGCTGCTCGTGGAGGCGCTCTAAAGCTGCTGCTGC	60		
1	GAATTCGGCAGAGGGGGTCTCGCTGCTCGTGGAGGCGCTCTAAAGCTGCTGCTGC	60		
61	GGCAGAGTTTGGAGGGGCGGCTTAAAGGTCATTTTCGTTGGGGGCTCCACAGGACCTT	120		
61	GGCAGAGTTTGGAGGGGCGGCTTAAAGGTCATTTTCGTTGGGGGCTCCACAGGACCTT	120		
121	CAGATCTCCGCTTAAAGTTCAGTTAAAGTTCGCGGAAAGCTGGGCGCAAGCGGTCACATGGCCA	180		
121	CAGATCTCCGCTTAAAGTTCAGTTAAAGTTCGCGGAAAGCTGGGCGCAAGCGGTCACATGGCCA	180		
181	CCCTTAACCTGGCGGAGCGCGGAGGCTCTGAGAAAGCCGGGACAGCCCGCTTTTCCCA	240		
181	CCCTTAACCTGGCGGAGCGCGGAGGCTCTGAGAAAGCCGGGACAGCCCGCTTTTCCCA	240		
241	GCCACCTCTAGAGTTTGGGACCAACAGAAAAGTGAAGTTCGCGGCTCTTCCAGAG	300		
241	GCCACCTCTAGAGTTTGGGACCAACAGAAAAGTGAAGTTCGCGGCTCTTCCAGAG	300		
301	CGTGGCGCAGCGCGCGCGCTGGGAGCAGAGTGGAGCGACCCCTTTACATAAAGATG	360		
301	CGTGGCGCAGCGCGCGCGCTGGGAGCAGAGTGGAGCGACCCCTTTACATAAAGATG	360		
361	AAAGCTGGGGTGGCTTACCTTTTGGGGGTCCTGCTGGAACTGCTGGGCTCGA	420		
361	AAAGCTGGGGTGGCTTACCTTTTGGGGGTCCTGCTGGAACTGCTGGGCTCGA	420		
421	AGGAGCCAAAGATCTTACACTGTGAGACTTGGAGGGCTGCTGGAGATTAAGTGGGAA	480		
421	AGGAGCCAAAGATCTTACACTGTGAGACTTGGAGGGCTGCTGGAGATTAAGTGGGAA	480		

QY	441	ATTCCCGCGGAGCCCAAGAACCATTCAGATGGGATCCCTCCGATCAATCCGAT	540
QY	441	ATTCCCGCGGAGCCCAAGAACCATTCAGATGGGATCCCTCCGATCAATCCGAT	540
Db	441	ATTCCCGCGGAGCCCAAGAACCATTCAGATGGGATCCCTCCGATCAATCCGAT	540
QY	541	GGCAGCCAGTCAGTGTGTGAGAGTACTTATGGCCGCTCAGAGGCCCACTTCACAGATTG	600
Db	541	GGCAGCCAGTCAGTGTGTGAGAGTACTTATGGCCGCTCAGAGGCCCACTTCACAGATTG	600
QY	601	CTTGAGGAGGTTGTGTGACCGAATGAAGAGAGTACGGGGAAACAGATTGACCTTACCCAC	660
Db	601	CTTGAGGAGGTTGTGTGACCGAATGAAGAGAGTACGGGGAAACAGATTGACCTTACCCAC	660
QY	661	CGCAAGAACTACGTACGCGCGTGCAGCCGGAATGGAATCCAGTAACCTAGACTTACAG	720
Db	661	CGCAAGAACTACGTACGCGCGTGCAGCCGGAATGGAATCCAGTAACCTAGACTTACAG	720
QY	721	GGCATCCGAAATGACTCAGATATACGCGGACCCCTCAAGTTTGGGTGTGAGAGCATTTGTG	780
Db	721	GGCATCCGAAATGACTCAGATATACGCGGACCCCTCAAGTTTGGGTGTGAGAGCATTTGTG	780
QY	781	GAAATATACGAGGATGAGCTTATCCGAATTTCTCTCACAAGAGGCTGACACAGTTAAAGAC	840
Db	781	GAAATATACGAGGATGAGCTTATCCGAATTTCTCTCACAAGAGGCTGACACAGTTAAAGAC	840
QY	841	AAACTTTGCAGTAAAGCGAGACAGATCATGTGTGACATCGCTGCACAGATCTCACAGATGAG	900
Db	841	AAACTTTGCAGTAAAGCGAGACAGATCATGTGTGACATCGCTGCACAGATCTCACAGATGAG	900
QY	901	CTATGAATCACTGTGAGCAGACGCTTACACCAACAGTATGAGAAACACCCCGAGAGGGGA	960
Db	901	CTATGAATCACTGTGAGCAGACGCTTACACCAACAGTATGAGAAACACCCCGAGAGGGGA	960
QY	961	AGATGGAGAGATGGCCTTTATATTATTCGTTTTATGGAATGAACTGAAAAAATCTCTTG	1020
Db	961	AGATGGAGAGATGGCCTTTATATTATTCGTTTTATGGAATGAACTGAAAAAATCTCTTG	1020
QY	1021	AAACCGAAGTAAAAAATTT	1069
Db	1021	AAACCGAAGTAAAAAATTT	1069
RESULT 2			
AAZ08294			
ID	AAZ08294	standard; DNA; 1069 BP.	
XX	AAZ08294;		
AC			
XX			
DT	07-FEB-2000	(first entry)	
XX			
DE	Mouse ortholog gene encoding Zs1g9 secretory protein.		
XX			
KW	Secretory protein-9; mouse Zs1g9; ortholog; overexpression; antagonist;		
KW	antibody; antisense nucleotide; tumour; treatment; receptor;		
KW	radio-label; polypeptide toxin; down-regulation; diagnostic;		
KW	therapeutic; probe; cancer; brain; liver; detection; stomach;		
KW	lymphoma; ds.		
XX			
OS	Mus musculus.		
XX			
XX			
FH	Key	Location/Qualifiers	
FT	CDS	358..906	
FT		/*tag= a	
FT		/product= "Mouse Zs1g9 secretory protein"	
FT		/note= "Overexpressed in tumours"	
FT	sig_peptide	358..417	
FT		/*tag= b	
FT	mat_peptide	418..903	
FT		/*tag= c	
FT		/label= Mature_Zs1g9_protein_variant-4	
XX			
XX	W09960405-A1.		



22-FEB-2000; 2000MO-US04414.  
 24-FEB-2000; 2000MO-US04914.  
 24-FEB-2000; 2000MO-US05004.  
 01-MAR-2000; 2000MO-US05004.  
 20-MAR-2000; 2000MO-US05377.  
 21-MAR-2000; 2000MO-US05377.  
 30-MAR-2000; 2000MO-US05377.  
 17-MAY-2000; 2000MO-US13705.  
 22-MAY-2000; 2000MO-US14042.  
 30-MAY-2000; 2000MO-US14941.  
 02-JUN-2000; 2000MO-US15264.  
 10-NOV-2000; 2000MO-US30873.  
 (GEM) GENENTECH INC.  
 Baker RP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W, Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S, Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WT, Zhang Z.  
 WPI. 2001-408281/43.  
 P-PSDB; AA012258.  
 Isolated, secretory and transmembrane PRO polypeptide used to detect other PRO polypeptides, link bioactive molecules to cells expressing PRO polypeptides, and detect the presence of mammalian tumours e.g. lung, breast, prostate, cervical.  
 Claim 3; Fig 173; 813pp; English.  
 AA012244-AS021518 encode for novel human secretory and transmembrane PRO polypeptides. The PRO polypeptides are useful to detect other PRO polypeptides, to link bioactive molecules to cells expressing PRO polypeptides, to modulate biological activities of cells expressing PRO polypeptides, and to detect the presence of mammalian lung, colon, breast, prostate, rectal, cervical or liver tumours by comparing PRO polypeptide expression in a cell sample to that in a control sample. Some of the 275 sequences are also useful to stimulate the release of tumor necrosis factor alpha (TNF-alpha) from human blood, the proliferation or differentiation of chondrocytes, the proliferation or gene expression in pericyte cells, the release of proteoglycans from cartilage, the proliferation of inner ear utricular supporting cells or of T-lymphocytes, or the release of a cytokine from peripheral blood monocytes (PBMCs), or the proliferation of endothelial cells. Some of the PRO polypeptides may modulate glucose or free fatty acid uptake by skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide to factor VIIa. The PRO polypeptides can be used in assays to identify molecules involved in binding interactions. The polynucleotides encoding PRO polypeptides can be used to generate probes, antisense RNA/DNA, transgenic or knock out animals and can be used in gene therapy.  
 Sequence 1210 BP; 261 A; 316 C; 379 G; 254 T; 0 other;  
 Query Match 64.6%; Score 691; DB 22; Length 1210;  
 Best Local Similarity 84.4%; Pred. No. 1.6e-169;  
 Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

435 -TAGTTTCGGAGCCACTGGGGCCAAAGTAGAGTCCACGGCTCTCCAGCCCTGGGCC 493  
 309 ACGGCGCGCGCCCTGGAGCAGAGTGGAGCGACCCCTGTAACTAATGATGAAGCGTG 368  
 494 ACGGCGCGCGCCCTGGAGCAGAGTGGAGCGACCCCTGTAACTAATGATGAAGCGTG 553  
 369 GGGTTGGCTAGCCCTAATTTTGGGGGCTCCTGGGAGTGGCTGGGCTGGAAGGCCA 428  
 554 GGGTTGGCTAGCCCTAATTTTGGGGGCTCCTGGGAGTGGCTGGGCTGGAAGGCCA 613  
 429 AGATCTACACTGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 488  
 614 GATCTCCACTGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 673  
 489 CGTGACCCCAAGAGACCATTCAGATGGATCCCTCCCAATCAATCAGATGGCCCA 548  
 674 GGTGACCCCAAGAGACCATTCAGATGGATCCCTCCCAATCAATCAGATGGCCCA 733  
 549 GTGAGTTGGAGTACCTTATGCCGCTCAGAGGCCCACTCAGAGTCTTGAGGA 608  
 734 GTGAGTTGGAGTACCTTATGCCGCTCAGAGGCCCACTCAGAGTCTTGAGGA 793  
 609 GGTGTGTACCAATGAAGAGTACGGGGAACAGATGACCTTCTACCCAGCCAGAA 668  
 794 GATGTGTACCAATGAAGAGTACGGGGAACAGATGACCTTCTACCCAGCCAGAA 853  
 669 CTACGTACCGCTGTGAGGAGCCGGAATGAGATTCAGATGACCTTCTACCCAGCC 728  
 854 CTACGTACCGCTGTGAGGAGCCGGAATGAGATTCAGATGACCTTCTACCCAGCC 913  
 729 AATGTACCTAGATATCAGGCGGACCTCAAGTTGGCTGAGAGCATTTGGAGAAATA 788  
 914 AATGTACCTAGATATCAGGCGGACCTCAAGTTGGCTGAGAGCATTTGGAGAAATA 973  
 789 CGAGATGAGCTTATGATTTCTTCCAGAGAGTGTCAACGTTAAAGAAACTTGG 848  
 974 CGAGATGAGCTTATGATTTCTTCCAGAGAGTGTCAACGTTAAAGAAACTTGG 1033  
 849 CAGTAAGCGAGATATCAGGCGGACCTCAAGTTGGCTGAGAGCATTTGGAGAAATA 908  
 1034 CAGTAAGCGAGATATCAGGCGGACCTCAAGTTGGCTGAGAGCATTTGGAGAAATA 1093  
 909 CACTGAGCAGACGCTTACACCAACGATGAGAAACCCCGAGAGGGAGATGCA 968  
 1094 CACTGAGCAGACGCTTACACCAACGATGAGAAACCCCGAGAGGGAGATGCA 1148  
 969 GCATGCGCTT--TATATGCTTTTATGAAATGAACTGAAAAA 1013  
 1149 GCATGCGCTTAT 1195  
 Db 1149 GCATGCGCTTAT  
 RESULT 4  
 AAX22112  
 ID AAX22112 standard: DNA; 1180 BP.  
 AAX22112;  
 18-MAY-1999 (first entry)  
 DE Human secreted protein gene 2 clone H2MB56.  
 XX Human; secreted protein; gene therapy; protein therapy; tissue; cancer;  
 XX tumour; neurodegenerative disorder; leukaemia; autoimmune disease;  
 XX developmental abnormality; foetal deficiency; Alzheimer's disease;  
 XX cognitive disorder; schizophrenia; immunological disorder; mood disorder;  
 XX immune deficiency disease; respiratory disorder; arthritis; skeletal;  
 XX haematopoietic disorder; neural; osteoporosis; metabolic disorders;  
 XX cardiovascular; endocrine; gastrointestinal; asthma; diagnosis; ds.  
 OS Homo sapiens.  
 XX  
 PN W09901020-A2.



PD 04-NOV-1999.  
 XX 27-APR-1999; 99WO-JP02226.  
 PF 28-APR-1998; 98JP-0119395.  
 PR (SAGA) SAGAMI CHEM RES CENT.  
 PA (PROT-) PROTEGENE INC.  
 XX Kato S, Kimura T;  
 PI WPI: 2000-023358/02.  
 DR P-PSDB: AAY52391.  
 XX  
 PT Human proteins with transmembrane domains, involved in control of cell  
 PT proliferation and differentiation, useful for treating e.g. cancer or  
 PT inflammation -  
 XX  
 PS Claim 4: Page 106-107; 114pp; English.  
 XX  
 CC This sequence represents the human cDNA clone HP10390  
 CC which encodes a 20 kb protein with one putative transmembrane  
 CC domain in the N-terminus. The cDNA was isolated from a human stomach  
 CC cancer cell line cDNA library. The protein has no homology with any  
 CC known protein. The protein may be used to raise specific antibodies, as  
 CC assay reagents, as diagnostic tissue markers, for the isolation of  
 CC cognate receptors, ligands and binding proteins, and as biologically  
 CC active agents. Nucleotides encoding the protein may be used as primers  
 CC and probes or antisense molecules, and in gene therapy. Cells transfected  
 CC with these nucleotides may be used to screen for agonists and antagonists  
 CC which are potentially useful therapeutically.  
 CC  
 SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;  
 Query Match 56.7%; Score 606.6; DB 21; Length 814;  
 Best Local Similarity 87.3%; Pred. No. 1.2e-147;  
 Matches 701; Conservative 0; Mismatches 94; Indels 8; Gaps 3;

OY 753 CCTCAAGTTTGGCTGTGAGAGCATTTGGAGAAATACGAGGATGACCTTATCGAATTCCT 812  
 DB 540 CCTCAAGTTTGGCTGTGAGAGCATTTGGAGAAATACGAGGATGACCTTATCGAATTCCT 599  
 OY 813 CTCACAGAGGCTGACAGGTTAAAGCAAACTTGCAGTAAGCGGACAGATCTAGTGA 872  
 DB 600 TTCACAGAGGCTGACAGGTTAAAGCAAACTTGCAGTAAGCGGACAGATCTTGTGA 659  
 OY 873 CCATGCCCTGCACAGATCTCACAGTATGATGAATCAGTACGACAGACGCTACACCA 932  
 DB 660 CCATGCCCTGCACAGATCTCACAGTATGATGAATCAGTACGACAGACGCTTGTGA 714  
 OY 933 AACGTATGAAACCCCGACAGAGGGAAGATGACGATTCCTT--TATATTACGTT 990  
 DB 715 GGCTTGATGATCACCACCGACAGAGGGAAGATGATGATTCCTTATATATATGTT 774  
 OY 991 TTTATGAAATGACGTGAAAAA 1013  
 DB 775 TTTACTGAAATTAATCTGAAAAA 797  
 RESULT 6  
 ABK52765  
 ID ABK52765 standard; cDNA; 814 BP.  
 XX  
 AC ABK52765;  
 XX  
 DT 15-AUG-2002 (first entry)  
 XX  
 DE cDNA encoding transmembrane protein 4, a cancer-linked protein.  
 XX  
 KW Expressed sequence tag; EST; human; cancer; anti-neoplastic;  
 KW cytotoxic T lymphocyte; chemotherapy; cytostatic; chromosome 12;  
 KW transmembrane protein 4; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 145..693  
 FT /tag= a  
 FT /product= "Transmembrane protein 4"  
 XX  
 PN W0200231198-A2.  
 XX  
 PD 18-APR-2002.  
 XX  
 PF 11-OCT-2001; 2001WO-US31607.  
 XX  
 PR 11-OCT-2000; 2000US-239294P.  
 PR 11-OCT-2000; 2000US-239297P.  
 PR 11-OCT-2000; 2000US-239605P.  
 PR 12-OCT-2000; 2000US-239802P.  
 PR 12-OCT-2000; 2000US-239805P.  
 PR 12-OCT-2000; 2000US-239806P.  
 PR 16-OCT-2000; 2000US-240622P.  
 PR 19-OCT-2000; 2000US-241682P.  
 PR 19-OCT-2000; 2000US-241723P.  
 PR 31-OCT-2000; 2000US-244932P.  
 XX  
 PA (AVAL-) AVALON PHARM.  
 XX  
 PI Young PE, Horrigan S, Weaver Z, Endress GA;  
 XX WPI: 2002-463271/49.  
 DR P-PSDB: AAV97063.  
 DR  
 PT Identifying modulators of a cancer-related gene to screen agents for  
 PT preventing or treating cancer comprises detecting a difference in the  
 PT expression of cancer-linked genes in the presence or absence of test  
 PT compounds -  
 XX  
 PS Claim 1: Page 43; 66pp; English.

XX The invention relates to modulators of a cancer-related genes. Also  
CC described are: (1) processes for identifying an anti-neoplastic agents  
CC comprising contacting a cell exhibiting neoplastic activity with a  
CC compound first identified as a cancer related gene modulator, and  
CC detecting a decrease in the neoplastic activity. (2) a process for  
CC determining the cancerous state of a cell by determining an increase in  
CC the level of expression of at least one gene, where an elevated  
CC expression relative to a known non-cancerous cell indicates a cancerous  
CC state or potentially cancerous state. The anti-neoplastic agent is  
CC useful for treating cancer or for protecting an animal against cancer.  
CC The immunogenic composition is also useful for treating cancer in an  
CC animal, where the composition elicits the production of cytotoxic T  
CC lymphocytes specific for the immunogenic composition. Preferably, the  
CC animal is a human. The cancer-linked genes and polypeptides are also  
CC useful as targets for cancer therapy or chemotherapy. The present  
CC sequence represents a cancer-linked gene located on chromosome 12,  
CC which encodes transmembrane protein 4.

SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Query Match 56.7%; Score 606.6; DB 24; Length 814;

Best Local Similarity 87.3%; Pred. No. 1.2e-147;  
Matches 701; Conservative 0; Mismatches 94; Indels 8; Gaps 3;

QY 213 AGAAGCCGGGACAGCCCGTTTCCAGCCAGCTGCTAGGGTTGGACCCACAGAAAC 272  
DB 1 AGAATCCCGGACAGCCCGTCTCCAGCCAGAGTG-TAGTTTCGGGAGCCACTGGGGCC 59  
QY 273 AAATGAGAGTCCGGGCTTTCCAGAGCCTGGGGCCAGGGGGCGGCGGCGGAGAGC 332  
DB 60 AAATGAGAGTCCAGGCTCTCCAGGCTGGGCGCCAGCGGGCGGCGGCGGAGAG 119  
QY 333 GTGAGAGGACCTGTACACATAAAGATGAAGGCTGGGTTGGCTAGCCCTACTTTGGG 392  
DB 120 GTGAGAGGACCTGTACACATAAAGATGAAGGCTGGGTTGGCTAGCCCTACTTTGGG 179  
QY 393 GGTCTGCTGGGAACTGCTGGGCTGAGAGAGCCAAATCTACACTGTGGAGCTTCAG 452  
DB 180 GGGCTGCTGGGAACTGCTGGGCTGAGAGAGCCAGATCTCAGTGTGAGAGATCAG 239  
QY 453 GGGCTGCTGGGAAATTAAGTGGGAAATTTGCCCGGCTGGAGCCCAAGAAACCATCA 512  
DB 240 GGGCTGCTGGGAAATTAAGTGGGAAATTTGCCCGGCTGGAGCCCAAGAAACCATCA 299  
QY 513 GATGGATCTCTCCGATCAATCCAGATGGAGCAGTGTGGAGGTTACCTTATGC 572  
DB 300 GATGGATCTCTCCGATCAATCCAGATGGAGCAGTGTGGAGGTTACCTTATGC 359  
QY 573 CCGCTCAGAGGCCACCTCAACAGATGCTTGAAGAGGTTGTGACCGAATGAAGAGTA 632  
DB 360 CCGCTCAGAGGCCACCTCAACAGATGCTTGAAGAGGTTGTGACCGAATGAAGAGTA 419  
QY 633 CGGGAACAGATGACCTCTACCCACCGCAGAACTAGCTACGCTCTGTAGCCGGA 692  
DB 420 TGGGGAACAGATGACCTCTACCCACCGCAGAACTAGCTACGCTCTGTAGCCGGA 479  
QY 693 TGGAGATCCAGTGAATAGTACAGGCGCATCCGAATGACTCAGATATCAGCGGAC 752  
DB 480 TGGAGATCCAGTGAATAGTACAGGCGCATCCGAATGACTCAGATATCAGCGGAC 539  
QY 753 CTTCAAGTTTGGCTGTAGAGCATTTGGAAGAAATACGAGATGACCTTATCGAATCTT 812  
DB 540 CTTCAAGTTTGGCTGTAGAGCATTTGGAAGAAATACGAGATGACCTTATCGAATCTT 599  
QY 813 CTTCAAGAGGCTGTAGAGCATTTAAAGCAAACTTTGCAATACCGGACATATCTATG 872  
DB 600 TTTCCGAGAGGCTGTAGAGCATTTAAAGCAAACTTTGCAATACCGGACATATCTATG 659  
QY 873 CCAATGCCCTGACAGATCTCAGATGAGTATGATCACTGAGAGCAAGCCCTACACA 932  
DB 660 CCAATGCCCTGACAGATCTCAGATGAGTATGATCACTGAGAGCAAGCCCTACACA 714

QY 933 AACGTGATGGAACACCCCGGAGGAGAGATGCGACATGCTTT--TATATTACGTT 990  
DB 715 GGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774  
QY 991 TTTATGGAATGAACTGAAATGAA 1013  
DB 775 TTTATGGAATTTAATGAAATGAA 797

RESULT 7

ID ABL41995  
AC ABL41995 standard; DNA; 814 BP.

AC ABL41995;

DT 11-JUN-2002 (first entry)

DE Nucleotide sequence of human polypeptide HP10390.

KW Antibody; antigen; transmembrane domain protein; HP10390; gene; ss.

XX Homo sapiens.

XX Key CDS 145..693

XX FT /tag- a /product- "HP10390"

PN W0200208416-A1.

PD 31-JAN-2002.

XX 24-JUL-2001; 2001WO-JP06371.

XX 24-JUL-2000; 2000JP-0222743.

PR 24-AUG-2000; 2000JP-0254407.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

PI Kato S, Nagata N, Fujimura N, Kobayashi M, Ito K, Ishizuka Y;

DR WPI; 2002-195877/25.

XX P-PSDB; ABB09717.

PT Antibody preparation by inoculation of an animal with a vector

PT expressing a fusion protein of an antigen on the C-terminal side of a

PT transmembrane domain for use as drugs, diagnostic reagents and

PT laboratory reagents

XX Example; Page 33-35; 45pp; Japanese.

XX The specification describes a method of antibody preparation. The

XX method comprises inoculating an animal with a vector expressing

XX a fusion protein having an antigen protein fused to the C-terminal

XX side (extracellular) of a transmembrane domain protein (the

XX N-terminal side of which is intracellular), and then isolating and

XX purifying the antibody from the animal. The antibodies can be used

XX as drugs, vaccines, diagnostic reagents and laboratory reagents. The

XX present sequence encodes a polypeptide, designated HP10390, which was

XX used in the course of the invention.

SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Query Match 56.7%; Score 606.6; DB 24; Length 814;

Best Local Similarity 87.3%; Pred. No. 1.2e-147;  
Matches 701; Conservative 0; Mismatches 94; Indels 8; Gaps 3;

QY 213 AGAAGCCGGGACAGCCCGTTTCCAGCCAGCTGCTAGGTTGGAGCCACAGAAAC 272.  
DB 1 AGAATCCCGGACAGCCCGTCTCCAGCAGAGTG-TAGTTTCGGGAGCCACTGGGGCC 59  
QY 273 AAATGAGAGTCCGGGCTTTCCAGAGCCTGGGCCAGGGCGGCGGCGGAGAGCAG 332  
DB 660 CCAATGCCCTGACAGATCTCAGATGAGTATGATCACTGAGAGCAAGCCCTACACA 714

DB 60 AAGTGAAGTCCAGGCTTCCAGCCGTTGGGCCACGGCGGCGCTGGAGCAGAG 119  
 QY 333 GTGAGACGACCGCTTTACATAAAGATGAAGGCTGGGGTGGCTAGCCCTACTTTTGGG 392  
 DB 120 GTGAGACGACCGCTTTACATAAAGATGAAGGCTGGGGTGGCTAGCCCTACTTTTGGG 179  
 QY 393 GGTCTGTGGTGAATAGAGTGGGAAATTTCCCGCTGGAGCCCAAGATCTACATTTGGAGCTTGGG 452  
 DB 180 GGCCTGTGGTGAATAGAGTGGGAAATTTCCCGCTGGAGCCCAAGATCTACATTTGGAGCTTGGG 239  
 QY 453 GGTCTGTGGTGAATAGAGTGGGAAATTTCCCGCTGGAGCCCAAGATCTACATTTGGAGCTTGGG 512  
 DB 240 GGTCTGTGGTGAATAGAGTGGGAAATTTCCCGCTGGAGCCCAAGATCTACATTTGGAGCTTGGG 299  
 QY 513 GATGGGATCTTCCGATCAATCCAGATGGCAGCCAGTCAATTTGGAGGATCTTATGC 572  
 DB 300 GATGGGATCTTCCGATCAATCCAGATGGCAGCCAGTCAATTTGGAGGATCTTATGC 359  
 QY 573 CGGCTCAGAGGCGCCACCTCAGAGTGTGTTGAGGAGGCTGTGACCCGATGAAGAGTA 632  
 DB 360 CGGCTCAGAGGCGCCACCTCAGAGTGTGTTGAGGAGGCTGTGACCCGATGAAGAGTA 419  
 QY 633 CGGGAACAGATTTGACCCCTTACCCAGCCAGAACTAGACGGCTGAGCCGGA 692  
 DB 420 TGGGGAGAGATTTGATCTTCCACCCATGCGAAGAACTAGTGTAGTGGCCGGA 479  
 QY 693 TGGAGATTCAGTGAATCTAGACTTACAGGGCATTCGAAATTTGACTGATATCAGCGGAC 752  
 DB 480 TGGAGATTCAGTGAATCTAGACTTACAGGGCATTCGAAATTTGACTGATATCAGCGGAC 539  
 QY 753 CCTCAAGTTGGTGTGAGAGATTTGAGAAATTCAGAGATGAGCTTATGAAATCTT 812  
 DB 540 CCTCAAGTTGGTGTGAGAGATTTGAGAAATTCAGAGATGAGCTTATGAAATCTT 599  
 QY 813 CTCAGAGAGGCTGACAACTTAAAGACAATTTGACATGAGCGGACAGATCTATGTA 872  
 DB 600 TCCCGAGAGGCTGACAACTTAAAGACAATTTGACATGAGCGGACAGATCTATGTA 659  
 QY 873 CCATGCCCCGACAGATCTCAGATGAGCTATGATCACTGAGACAGACCTTACCA 932  
 DB 660 CCATGCCCCGACAGATCTCAGATGAGCTATGATCACTGAGACAGACCTTACCA 714  
 QY 933 AACGTATGGAACACCCCGAGGAGGAAAGATGGCAGCTTGCCTT--TATATTACGT 990  
 DB 715 GGTGTGATGATCACCACCCAGAGGAGAAATGTGGCAATGCTTTATATTATGTT 774  
 QY 991 TTTATGGAATGAACTGAAAAA 1013  
 DB 775 TTTACTGAAATTTACTGAAAAA 797  
 RESULT 8  
 ID ABR09772 standard; cDNA; 814 BP.  
 AC ABR09772;  
 DT 14-MAR-2002 (first entry)  
 DE Human ovarian tumour protein encoding cDNA #305.  
 KW Human; ovarian tumour protein; cancer; cytostatic; immunostimulant; ss;  
 KW gene therapy; CD4+ T cell; CD8+ T cell; PCR primer.  
 OS Homo sapiens.  
 XX MO200190154-A2.  
 XX 29-NOV-2001.  
 XX 23-MAY-2001; 2001MO-US16895.  
 XX 24-MAY-2000; 2000US-207107P.  
 PR

PR 13-JUN-2000; 2000US-211457P.  
 PR 21-JUN-2000; 2000US-213673P.  
 PR 03-AUG-2000; 2000US-223288P.  
 PR 01-MAR-2001; 2001US-272790P.  
 PA (COR-) CORIXA CORP.  
 PI Xu J, Mitcham JL, Harlocker SL, Dillon DC, Secrist H, Iodes MJ;  
 PI Algate PA, Fling SP, Mannion J, Benson DR, Carter D;  
 DR WPI; 2002-097641/13.  
 XX  
 PT New isolated polynucleotide encoding polypeptide comprising portion of  
 PT ovarian tumour protein, useful for detection, diagnosis and therapy of  
 PT human ovarian cancer  
 PS  
 XX  
 CC Claim 1; Page 253-254; 285pp; English.  
 CC The invention relates to an isolated polynucleotide encoding a  
 CC polypeptide comprising a portion of an ovarian tumour protein. The  
 CC sequences of the invention are useful for stimulating an immune response  
 CC and for treating ovarian cancer in a patient. An antigen presenting cell  
 CC that expresses the sequences is useful for treating ovarian cancer by  
 CC incubating CD4+ and/or CD8+ T cells isolated from a patient. The T cells  
 CC can then be proliferated and administered to the patient to inhibit the  
 CC development of cancer. The DNA sequences are useful as probes or primers  
 CC for nucleic acid hybridisation, to direct expression of a polypeptide in  
 CC appropriate host cells. Detecting the presence of a cancer in a patient  
 CC involves obtaining a biological sample from the patient, contacting the  
 CC biological sample with an agent that binds to the protein, detecting the  
 CC amount of protein that binds to the agent, comparing the amount of  
 CC protein to a predetermined cut-off value and determining the presence of  
 CC cancer. Sequences ABR09464-ABR09802 represent PCR primers and cDNA  
 CC molecules encoding ovarian tumour proteins of the invention.  
 SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;  
 Query Match 56.7%; Score 606.6; DB 24; Length 814;  
 Best Local Similarity 87.3%; Pred. No. 1.2e-147;  
 Matches 701; Conservative 0; Mismatches 94; Indels 8; Gaps 3;  
 QY 213 AAGAGCCGGGAGACGCCGCTTTCCAGCAGCTGTAGGGTGGGAGCCACAGAAAC 272  
 DB 1 AGAATCCGGAGACGCCCTGCTCCTCAGCAGAGTG-TGTTTGGGAGCCATGGGCC 59  
 QY 273 AAGTGAAGTCCGCTGCTTCCAGAGCCTGGGCCAGCGCGCGCTGGAGCAGAG 332  
 DB 60 AAGTGAAGTCCAGCGGCTTCCAGCGCTTGGGCCAGCGCGCGCTGGAGCAGAG 119  
 QY 333 GTGAGACGACCGCTTTACATAAAGATGAAGGCTGGGTTGGCTAGCCCTACTTTTGGG 392  
 DB 120 GTGAGACGACCGCTTTACATAAAGATGAAGGCTGGGTTGGCTAGCCCTACTTTTGGG 179  
 QY 393 GGTCTGTGGTGAATAGAGTGGGAAATTTCCCGCTGGAGCCCAAGATCTACATTTGGAGCTTGGG 452  
 DB 180 GGCCTGTGGTGAATAGAGTGGGAAATTTCCCGCTGGAGCCCAAGATCTACATTTGGAGCTTGGG 239  
 QY 453 GGTCTGTGGTGAATAGAGTGGGAAATTTCCCGCTGGAGCCCAAGATCTACATTTGGAGCTTGGG 512  
 DB 240 GGTCTGTGGTGAATAGAGTGGGAAATTTCCCGCTGGAGCCCAAGATCTACATTTGGAGCTTGGG 299  
 QY 513 GATGGGATCTTCCGATCAATCCAGATGGCAGCCAGTCAATTTGGAGGATCTTATGC 572  
 DB 300 GATGGGATCTTCCGATCAATCCAGATGGCAGCCAGTCAATTTGGAGGATCTTATGC 359  
 QY 573 CGGCTCAGAGGCGCCACCTCAGAGTGTGTTGAGGAGGCTGTGACCCGATGAAGAGTA 632  
 DB 360 CGGCTCAGAGGCGCCACCTCAGAGTGTGTTGAGGAGGCTGTGACCCGATGAAGAGTA 419  
 QY 633 CGGGAACAGATTTGACCCCTTACCCAGCCAGAACTAGACGGCTGAGCCGGA 692  
 DB 420 TGGGGAGAGATTTGATCTTCCACCCATGCGAAGAACTAGTGTAGTGGCCGGA 479



QY 693 TGGAGAAATCCAGTACTAGACTTACAGGCGATCCGAATTGACTCAGATATACGCGGCAC 752  
 Db 480 TGGAGAAATCCAGTACTAGACTTACAGGCGATCCGAATTGACTCAGATATACGCGGCAC 539  
 QY 753 CCTCAAGTTTGGCTGTAGAGCATTTGGAGAAATAGAGGATGAGCTTATCGAATCTT 812  
 Db 540 CCTCAAGTTTGGCTGTAGAGCATTTGGAGAAATAGAGGATGAGCTTATCGAATCTT 599  
 QY 813 CTCCAGAGAGGCTGACACCTTAAAGACAACCTTGGAGTACGCGAGCATCTATGTGA 872  
 Db 600 TTCCCGAGAGGCTGACACCTTAAAGACAACCTTGGAGTACGCGAGCATCTATGTGA 659  
 QY 873 CCATGCGCTGCAGATATCGATGATGATGATGATGATGATGATGATGATGATGATGAT 932  
 Db 660 CCATGCGCTGCAGATATCGATGATGATGATGATGATGATGATGATGATGATGATGAT 714  
 QY 933 AACGTATGAGACACCCCGAGAGGAGAGATGAGCATTTGCCCTT--TATTTACGTT 990  
 Db 715 GGGTTGATGATCACCCTCCAGAGGAGGAAATGTTGGCATGCTTTTATATATTTATGTT 774  
 QY 991 TTTATGAAATGAACTGAAATAA 1013  
 Db 775 TTTACTGAAATTTACTGAAATAA 797

RESULT 9  
 AAX06969  
 ID AAX06969 standard; cDNA; 806 BP.  
 XX AAX06969;  
 AC AAX06969;  
 XX  
 DT 10-MAY-1999 (first entry)  
 XX  
 DE Human secretory peptide-9 (Zs19) variant cDNA.  
 XX  
 KW Secretory peptide-9; Zs19; human; tumour marker; cancer; therapy;  
 diagnosis; growth enhancer; variant; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 104..652  
 FT /tag- a  
 FT sig-peptide 104..163  
 FT /tag- b  
 FT mat-peptide 164..649  
 FT /tag- c  
 XX  
 PN WO9901554-A1.  
 XX  
 PD 14-JAN-1999.  
 XX  
 PF 02-JUL-1998; 98WO-US13859.  
 XX  
 PR 17-JUN-1998; 98US-0099005.  
 PR 03-JUL-1997; 97US-0051704.  
 PR 13-JUL-1997; 97US-0888088.  
 PR 19-MAY-1998; 98US-0081338.  
 PR 19-MAY-1998; 98US-0083883.  
 PR 17-JUN-1998; 98US-0089899.  
 XX  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 PI Jaspers SR, Jelinek LJ, Sheppard PO, Whitmore TE;  
 DR WPI: 1999-106055/09.  
 DR P-PSDB; AAM88474.  
 XX  
 PT New mammalian secretory peptide-9 (Zs19) - used as a growth  
 PT enhancer for placenta, liver and heart, and as an indicator of  
 PT cancer  
 PS Claim 2; Page 73-74; 85pp; English.

XX This cDNA clone encodes human secretory peptide-9, or Zs19,  
 CC variant (see AAM88474). Zs19 (see also AAM88469) is overexpressed in  
 CC human brain, liver, lung, oesophageal, stomach, colon, rectal,  
 CC thyroid and lymphoma tumors. Thus, Zs19 can be used as an  
 CC indicator for cancer. Zs19 cDNA was discovered in a placenta  
 CC clone from a full-term pregnancy cDNA library which contained an  
 CC expressed sequence tag (see AAX06971). The invention provides an  
 CC polynucleotide (see AAX06968-70) encoding Zs19 polypeptides (see  
 CC AAM88469-77) including mature polypeptides, other processed forms,  
 CC variants and mouse orthologues. The Zs19 gene, or probes derived  
 CC from it, can be used to determine if Zs19 is present on chromosome  
 CC 12, and if a mutation has occurred. Antibodies raised against  
 CC Zs19 can be used as diagnostic agents to determine the presence of  
 CC Zs19, and thus the presence of cancer. They can also be labelled  
 CC with radioisotopes or fused with toxins and used to treat tumours  
 CC which overexpress Zs19. Antisense nucleotides derived from Zs19  
 CC cDNA can also be used to inhibit the growth of tumour cells. Zs19  
 CC proteins can be used to enhance the growth or development of the  
 CC placenta, heart or liver.

Sequence 806 BP; 229 A; 183 C; 232 G; 160 T; 2 other;

Query Match 56.1%; Score 599.6; DB 20; Length 806;  
 Best Local Similarity 87.7%; Pred. No. 7.7e-146;  
 Matches 690; Conservative 1; Mismatches 86; Indels 10; Gaps 3;

QY 272 CAAAGTGAAGATCCGGCTCTTTCAGAGCCCTGGCCAGCGCGCGCGGTGGAGCAGA 331  
 Db 18 CAAAGTGAAGATCCGGCTCTTTCAGAGCCCTGGCCAGCGCGCGCGGTGGAGCAGA 77  
 QY 332 GGTGAGACGACCCCTGTTACACTAAAGATGAAGCTGGGTGGCTTACCTTCTTGG 391  
 Db 78 GGTGAGACGACCCCTGTTACACTAAAGATGAAGCTGGGTGGCTTACCTTCTTGG 137  
 QY 392 GGGTCTGTGGGAACTGCTGGCTGGAAGAGCCAAATCTACACTGTGGAGCTTGA 451  
 Db 138 GGGCTGTGGGAACTGCTGGCTGGAAGAGCCAAATCTACACTGTGGAGCTTGA 197  
 QY 452 GGGCTGTGGGAACTGCTGGCTGGAAGAGCCAAATCTACACTGTGGAGCTTGA 511  
 Db 198 GGGCTGTGGGAACTGCTGGCTGGAAGAGCCAAATCTACACTGTGGAGCTTGA 257  
 QY 512 AGATGGATCTTCGATCAATCCAGATGCGAGCCAGTGTGGAGATGATCTTATG 571  
 Db 258 AGATGGATCTTCGATCAATCCAGATGCGAGCCAGTGTGGAGATGATCTTATG 317  
 QY 572 CCGGCTCAGAGGCCCACTCAGAGATCTTGGAGGAGTGTGACCGAATGAAGAGT 631  
 Db 318 CCGGCTCAGAGGCCCACTCAGAGATCTTGGAGGAGTGTGACCGAATGAAGAGT 377  
 QY 632 ACGGGAGACAGATTGACCTTTCACCCAGCAAGAACTACGCTGACCGCGA 691  
 Db 378 ACGGGAGACAGATTGACCTTTCACCCAGCAAGAACTACGCTGACCGCGA 437  
 QY 692 ATGGAGAAATCCAGTGAATCTTACAGGAGCATCCGAATCTACATATCAGCGCA 751  
 Db 438 ATGGAGAAATCCAGTGAATCTTACAGGAGCATCCGAATCTACATATCAGCGCA 497  
 QY 752 CCTCAAGTTTGGCTGTAGAGCATTTGGAGAAATAGAGGATGAGCTTATCGAATCT 811  
 Db 498 CCTCAAGTTTGGCTGTAGAGCATTTGGAGAAATAGAGGATGAGCTTATCGAATCT 557  
 QY 812 TCTCCAGAGAGGCTGACACCTTAAAGACAACCTTGGAGTACGCGAGCATGTATGTG 871  
 Db 558 TCTCCAGAGAGGCTGACACCTTAAAGACAACCTTGGAGTACGCGAGCATGTATGTG 617  
 QY 872 AACATGCCCTGACAGATCTCAGATGAGCTATGATATCTGAGACAGACCTTACACC 931  
 Db 618 AACATGCCCTGACAGATCTCAGATGAGCTATGATATCTGAGACAGACCTTACACC 672  
 QY 932 AAAGTGTGAAACACCCCGAGAGGAGAGATGAGCATTTGCCCTT--TATATTACGT 989



```

XX  Human secreted protein encoding cDNA #25.
DE
XX
XX  Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;
KW  diagnostic; gene therapy; chromosome mapping; secretion vector; ss.
OS
XX  Homo sapiens.
XX
XX  WO925825-A2.
PN
XX
XX  27-MAY-1999.
PD
XX
XX  13-NOV-1998; 98WO-IB01862.
PF
XX
XX  04-SEP-1998; 98US-0099273.
PR  13-NOV-1997; 97US-0066677.
PR  17-DEC-1997; 97US-0069957.
PR  09-FEB-1998; 98US-0074121.
PR  13-APR-1998; 98US-0081563.
PR  10-AUG-1998; 98US-0096116.
XX
XX  (GEST ) GENSET.
PA
XX
XX  Bougueleret L, Duclert A, Dumas Milne Edwards J;
PI  WPI: 1999-347472/29.
XX  P-PSDB; AAY36153.
DR
XX
XX  Extended cDNAs encoding secreted proteins
PT
XX
XX  Example 28; Page 194-195; 307Pp; English.
PS
XX
XX  AAY97813-X97906 represent extended cDNA's which encode novel human
CC  secreted proteins (see AAY36129-Y36222) and which have cytosolic,
CC  thrombotic and osteopathic activity. The extended cDNAs can be used to
CC  express secreted proteins or parts of them or to obtain antibodies
CC  capable of binding to the secreted proteins. They may also be used in
CC  diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC  Uses also include design of expression vectors and secretion vectors.
XX
XX
XX  Sequence 832 BP; 224 A; 194 C; 243 G; 168 T; 3 other;
SQ

```

Query Match 52.88; Score 564.6; DB 20; Length 832;

Best Local Similarity 85.98; Pred. No. 9.7e-137;

Matches 722; Conservative 0; Mismatches 104; Indels 15; Gaps 8;

```

QY 211 GGAGAACCCGGGACAGCCCGCTTTTCCAGCCAGCTGTGAGGTGGGACCCAGAGAA 270
DB 2  GGAGAAATCCGGAGACGCTGCTCCGACAGGTG-TAGTTTCGGAGCCACTGGGG 60
QY 271 ACAAGTGAAGTCCGGCTCTCTTCCAGAGCTGGGCCAGCGGCGGCGCTGGAGCAG 330
DB 61 CCAAGTGAAGTCCAGCGCTTCCAGCGCTTGGGCCAGCGGCGGCGCTGGAGCAG 120
QY 331 AGGTGAGCGACCCGTTTACATTAAGATGAAGAGCTGGGCTTGGCTAGCCCTTACTTTG 390
DB 121 AGGTGAGCGACCCGTTTACATTAAGATGAAGAGCTGGGCTTGGCTAGCCCTTACTTTG 180
QY 391 GGGGCTCTGCTGGAACTGCTGGCTCGAAGAGCGCAAGATGTACAGTGGAGCTTGC 450
DB 181 GGGGCTCTGCTGGAACTGCTGGCTCGAAGAGCGCAAGATGTACAGTGGAGCTTGC 240
QY 451 AGGGCTCTGCTGGATG-AATTAGTGGGAATTTGCCCGCTGGAGCCCAAGAACAT 509
DB 241 AGGGCTCTGCTGGATGAATTAAGTGAATTTGCCCGCTGGAGCCCAAGAACAT 300
QY 510 TCAGATGGGATCTTCCGATCAATCCAGATGGAGCCGCTAGTGTGGAGAGTACTTA 569
DB 301 TCAGATGGGATCTTCCGATCAATCCAGATGGAGCCGCTAGTGTGGAGAGTACTTA 360
QY 570 TGCCTGCTAGAGGCCACCTCAGAGAGTGTCTTGAAGAGGTGTGAACCAAGTGAAGA 629
DB 361 TGCCTGCTAGAGGCCACCTCAGAGAGTGTCTTGAAGAGGTGTGAACCAAGTGAAGA 420

```

```

QY 630 GTACGGGAACAGATTGACCTTCTACCCACCGCAAGAACTAGCTAGCGCTGTGACCG 689
DB 421 GTATGGGGAACAGATTGATCTCTTCCACCCATCGCAAGAACTAGCTAGCTGTGACCG 480
QY 690 GAATGAGAAATCCAGTGAAGTACTTACAGGCAATCCGAATTTAGTACATATACCGG 749
DB 481 GATGAGAAATCCAGTGAAGTACTTACAGGCAATCCGAATTTAGTACATATACCGG 540
QY 750 CACCTCAAG-TTTCGCTGTGAGAGCATTTGGAAGAATACGAGAGTACTTATGCAAT 808
DB 541 CACCTCAAGBTTCGCTGTGAGAGCATTTGGAAGAATACGAGAGTACTTATGCAAT 600
QY 809 TCTTCTCCAGAGAGCGTGCACACTTAAAGCAAACTTTCAGTAAAGCGAGACATCTAT 868
DB 601 TCTTCTCCAGAGAGCGTGCACACTTAAAGCAAACTTTCAGTAAAGCGAGACATCTT 660
QY 869 GTGACCATGCGCTGCACAGATC-TCAGGATGAGTATGATACACTGAGAGCAGACCTTA 927
DB 661 GTGACCATGCGCTGCACAGATC-TCAGGATGAGTATGATACACTGAGAGCAGACCTTA 717
QY 928 CACCAAACTGATGGAACACCCCGCAGG-AGGGGAAGATGCGACATTCCTTTATAT-T 985
DB 718 -ACTGCTTGTATGATGATCACCCCGAGNAAGGAAATGCTGCAATGCTTTATATAT 775
QY 986 ACGTTTATGAAATGAGTGAAGAAAACTTTGAACCGAAGTAAAAA 1045
DB 776 TATGTTTACTGAATTAATTAAGTAAAAATA---TGAACCAAAAGTSCAAAAA 831
QY 1046 A 1046
DB 832 A 832

```

RESULT 12

ID AAX97884 standard; cDNA; 714 BP.

AAX97884

AAX97884;

AC AAX97884;

DE 23-SEP-1999 (first entry)

KW Human secreted protein encoding cDNA #72.

KW Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;

OS diagnostic; gene therapy; chromosome mapping; secretion vector; ss.

OS Homo sapiens.

PN WO925825-A2.

PD 27-MAY-1999.

PF 13-NOV-1998; 98WO-IB01862.

PR 04-SEP-1998; 98US-0099273.

PR 13-NOV-1997; 97US-0066677.

PR 17-DEC-1997; 97US-0069957.

PR 09-FEB-1998; 98US-0074121.

PR 13-APR-1998; 98US-0081563.

PR 10-AUG-1998; 98US-0096116.

PA (GEST ) GENSET.

PI Bougueleret L, Duclert A, Dumas Milne Edwards J;

DR WPI: 1999-347472/29.

DR P-PSDB; AAY36200.

PT Extended cDNAs encoding secreted proteins

PS Claim 1; Page 272; 307Pp; English.

CC AAX97813-X97906 represent extended CDNA's which encode novel human  
CC secreted proteins (see AAY36129-136222) and which have cytostatic,  
CC thrombotic and osteopathic activity. The extended CDNA's can be used to  
CC express secreted proteins or parts of them or to obtain antibodies  
CC capable of binding to the secreted proteins. They may also be used in  
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC Uses also include design of expression vectors and secretion vectors.  
XX

Sequence 714 BP; 203 A; 158 C; 202 G; 151 T; 0 other;

Query Match 51.9%; Score 555.2; DB 20; Length 714;  
Best Local Similarity 88.5%; Pred. No. 2.5e-134;  
Matches 639; Conservative 0; Mismatches 73; Indels 10; Gaps 3;

OY 336 ACCAGAGTGAAGCACCCTGTTACATAAAGATGAAAGGCTGGGTTGGTACGCCCTAC 385  
DB 1 AGCAGAGGTGAGCAGCACCCTTACGCTTAAGATGAAAGGCTGGGTTGGTACGCCCTAC 60  
OY 386 TTTGGGGGCTGCTGGGAACTGCTGGGCTGAAAGGAGGCAAGATCTACACTGTGGAG 445  
DB 61 TTTGGGGGCTGCTGGGAACTGCTGGGCTGAAAGGAGGCAAGATCTACACTGTGGAG 120  
OY 446 CTTCGAGGCTCTGCTGGATGATTAAGATGAGAAATTCCTGGGAGCCCAAGAGA 505  
DB 121 CATGAGGCTCTGCTGGATGATTAAGATGAGAAATTCCTGGGAGCCCAAGAGA 180  
OY 506 CCATTCAGATGGGATCTCTCCGATCAATCCAGATGGCAGCCAGTCTGTTGGAGTAC 565  
DB 181 CCATTCAGATGGGATCTCTCCGATCAATCCAGATGGCAGCCAGTCTGTTGGAGTAC 240  
OY 566 CTATAGCCGCTCAGAGGCCCTCAGACAGTCTGAGAGAGGCTGTGACCGAATTA 625  
DB 241 CTATAGCCGCTCAGAGGCCCTCAGACAGTCTGAGAGAGGCTGTGACCGAATTA 300  
OY 626 AGGAGTACGGGAAAGATGATGACCTTACCCAGCCGCAAGAACTACATGACCTGCTGA 685  
DB 301 AGGAGTACGGGAAAGATGATGACCTTACCCAGCCAGCAAGAACTACATGACCTGCTGA 360  
OY 686 GCCGGAATGGAGAAATCCAGTACATGATGATGATGATGATGATGATGATGATGATGAT 745  
DB 361 GCCGGAATGGAGAAATCCAGTACATGATGATGATGATGATGATGATGATGATGATGAT 420  
OY 746 GCGGACACCTCAAGTTGGTGTGAGAGCATGTTGGAAGAAATACGAGATGATGATGATG 805  
DB 421 GCGGACACCTCAAGTTGGTGTGAGAGCATGTTGGAAGAAATACGAGATGATGATGATG 480  
OY 806 AATTCCTCCAGAGAGGCTGCAACGTTAAAGCAAACTTTGCAAGTAAAGCCGAGATG 865  
DB 481 AATTCCTCCAGAGAGGCTGCAACGTTAAAGCAAACTTTGCAAGTAAAGCCGAGATG 540  
OY 866 TATGTCAGACATGCTGCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 925  
DB 541 TATGTCAGACATGCTGCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 599  
OY 926 TACACCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 983  
DB 600 ---ACTGGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 655  
OY 984 TTACGTTTTTATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1043  
DB 656 TTATGTTTTTACTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 712  
OY 1044 AA 1045  
DB 713 AA 714

RESULT 13  
AAX43525  
ID AAX43525 standard; CDNA: 592 BP.  
XX  
AC AAX43525;  
XX

DT 21-AUG-2000 (first entry)  
XX  
DE Mouse secreted expressed sequence tag SEQ ID NO:100.  
XX

KW Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;  
KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
KW immunomodulatory; haematopoietic; chemokine; analgesic; haemostatic;  
KW thrombolytic; antiinflammatory; cytoskeletal; antibacterial; antifungal;  
KW antiviral; antidiabetic; antiaesthetic; vulnery; antiparkinsonian;  
KW antitumor; osteopathic; neuroprotective; nootropic; antiproliferative;  
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
KW central nervous system disorder; Alzheimer's disease; stroke;  
KW Parkinson's disease; Huntington's disease; coagulation disorder;  
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
KW tumor; infection; depression; psoriasis; ss.

OS Mus musculus.  
XX  
PN WO20021991-A1.  
XX

PD 20-APR-2000.  
XX

PF 15-OCT-1999; 99WO-US24206.  
XX

PR 15-OCT-1998; 98US-0104436.  
XX

PA (GENY) GENETICS INST INC.  
XX

PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racle LA, Evans C;  
PI Merberg D, Treacy M, Bowman MR;  
XX

DR WPI: 2000-317938/27.  
XX

PT Isolated polynucleotides, and encoded proteins, comprising secreted  
PT expressed sequence tags (SESTs), useful for treating various disorders  
PT such as autoimmune, infectious, and central nervous system disorders -  
XX  
PS Claim 1; Page 228-229; 803pp; English.

CC AAX43426 to AAX45925 represent specifically claimed secreted expressed  
CC sequence tags (SESTs), isolated from human, mouse, chicken and rat  
CC tissue sources. The SESTs can have a range of activities depending on  
CC the tissues they were isolated from. The activities include:  
CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
CC chemokine; analgesic; haemostatic; thrombolytic; antiinflammatory;  
CC cytoskeletal; antibacterial; antifungal; antiviral; antidiabetic;  
CC antiaesthetic; vulnery; antitumor; osteopathic; neuroprotective;  
CC nootropic; antiparkinsonian; antiproliferative; cerebroprotective;  
CC anticonvulsant; and antidepressant. The SESTs can be used for gene  
CC therapy and in vaccines. The SESTs are useful as probes for the  
CC identification and isolation of full-length CDNA's and genomic DNA  
CC molecules which correspond to the SESTs. Proteins encoded by the SESTs  
CC are useful in assays for determining biological activity and raising  
CC antibodies. They may be useful for treatment of autoimmune disorders  
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
CC osteoporosis, osteoarthritis, central nervous system disorders  
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
CC disease), tumors, bacterial, fungal or viral infections, depression and  
CC psoriasis. AAX45926 to AAX45931 represent linker variants which are given  
CC in the exemplification of the present invention.

SO Sequence 592 BP; 159 A; 139 C; 167 G; 123 T; 4 other;

Query Match 48.0%; Score 513.4; DB 21; Length 592;  
Best Local Similarity 97.7%; Pred. No. 1.7e-123;  
Matches 550; Conservative 0; Mismatches 10; Indels 3; Gaps 3;

OY 329 AGAGTGAAGCACCCTGTTACATAAAGATGAAAGGCTGGGTTGGCTAG-CCCTACTT 387

```

Db 15 AGGCGTAGAGGACCGCTTACACTAAAGATGAAAGGNTGGGCTGCTGCTACTT 74
Qy 388 TTGGGGGCTCTGCTGGGAGTGCCTGGGCTCAAGAGAGCCAGATGTACCTGTGAGCT 447
Db 75 TTGGGGGCTCTGCTGGGAGTGCCTGGGCTCAAGAGAGCCAGATGTACCTGTGAGCT 134
Qy 448 TTGAGGGCTCTGCTGGGAGTGCCTGGGCTCAAGAGAGCCAGATGTACCTGTGAGCT 507
Db 135 TGCAGGGCTCTGCTGGGAGTGCCTGGGCTCAAGAGAGCCAGATGTACCTGTGAGCT 194
Qy 508 ATTCAAGTGGGAGTCTTCCGATCAATCCAGATGGGAGGAGGAGTGTGAGAGTACT 567
Db 195 ATTCAAGTGGGAGTCTTCCGATCAATCCAGATGGGAGGAGGAGTGTGAGAGTACT 254
Qy 568 TATGCGGCTCTGAGAGGAGGAGGAGGAGTGTGAGAGGAGTGTGAGAGGAGTGTGAG 627
Db 255 TATGCGGCTCTGAGAGGAGGAGGAGGAGTGTGAGAGGAGTGTGAGAGGAGTGTGAG 314
Qy 628 GAGTACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 687
Db 315 GAGTACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 372
Qy 688 CCGAATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 747
Db 373 CCGAATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 432
Qy 748 GGCACCTCAAGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 807
Db 433 GGCACCTCAAGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 492
Qy 808 TTCTTCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 867
Db 493 TTCTTCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 552
Qy 868 TGTGACCATGCGCTGACAGATG 890
Db 553 TGTGACCATGCGCTGACAGATG 575

RESULT 14
ABK36007
ID ABK36007 standard; cDNA; 657 BP.
XX
AC ABK36007;
XX
DT 08-MAY-2002 (first entry)
XX
DE cDNA sequence #398 encoding novel human secreted protein.
XX
KM Human secreted protein: hyperproliferative disorder; autoimmune disorder;
KM immune deficiency disorder; blood disorder; inflammatory disorder;
KM infectious disorder; allergic condition; neurodegenerative disorder;
KM liver fibrosis; coagulation disorder; gene therapy; antimicrobial;
KM tumour; cancer; hepatotropic; immunosuppressive; antineutrotic; gene; ss.
XX
OS Homo sapiens.
XX
PN MO200177289-A2.
XX
PD 18-OCT-2001.
XX
PF 29-MAR-2001; 2001WO-US10232.
XX
PR 06-APR-2000; 2000US-195605P.
XX
PA (GENY) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, Lavalie ER, Collins-Racie LA, Evans C,
PI Metberg D, Treacy M, Agostino MJ, Bowman MR, Spaulding V, Wong GG;
PI Clark HF, Fechtel K, Howes SH, Resnick RJ, Gulukota K, Graham JR;
XX
WPI: 2002-179322/23.

```

XX Six hundred and twenty three polynucleotides derived from a variety of  
PT human tissue sources which encode secreted proteins, useful for  
PT treating immune deficiencies and disorders such as autoimmune disorders  
XX  
PS Claim 1; Page 297; 393pp; English.

CC The present invention relates to the isolation of novel cDNA sequences  
CC which encode human secreted proteins. The cDNA sequences have been  
CC derived from a variety of human tissues. The invention also provides  
CC a method for producing proteins from these polynucleotide sequences.  
CC The proteins are useful for identifying compounds that modulate their  
CC activity and production. The sequences of the invention are  
CC useful for treating diseases such as hyperproliferative disorders  
CC (e.g. cancer), immune deficiency disorders (e.g. severe combined  
CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple  
CC sclerosis), blood disorders (e.g. thrombocytopaenia), inflammatory  
CC disorders (e.g. arthritis), infectious disorders (e.g. hepatitis),  
CC allergic conditions (e.g. asthma), neurodegenerative disorders (e.g.  
CC Alzheimer's disease), liver fibrosis, coagulation disorders (e.g.  
CC hemophilia), and tumours. The polynucleotide sequences of the  
CC invention are also useful in gene therapy. ABK3610-ABK36232 represent  
CC the cDNA sequences of the invention that encode for novel human  
CC secreted proteins.

SQ Sequence 657 BP: 177 A; 149 C; 185 G; 146 T; 0 other;

Query Match 47.1%; Score 503.2; DB 24; Length 657;  
Best Local Similarity 88.4%; Pred. No. 86-121;

Matches 571; Conservative 0; Mismatches 68; Indels 7; Gaps 2;

```

Qy 370 GGTGGCTGCTCTACTTTGGGGGCTCTGCTGGGAGTGCCTGGGCTCGAAGAGCCAA 429
Db 1 GGTGGCTGCTCTCTCTGCTGGGAGTGCCTGGGAGTGCCTGGGCTCGAAGAGCCAA 60
Qy 430 GATCTACACTGTGAGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGG 489
Db 61 GATCTCACTGTGAGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGG 120
Qy 490 GTGGAGCCCAAGAGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGG 549
Db 121 GTGGAGCCCAAGAGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGG 180
Qy 550 TCAGTGTGGAGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAG 609
Db 181 TCAGTGTGGAGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAG 240
Qy 610 GTCGTGACCGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAGTGC 669
Db 241 ATATGTGACCGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAGTGC 300
Qy 670 TACGTACCGTGTGAGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGG 729
Db 301 TACGTACCGTGTGAGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGG 360
Qy 730 ATTGACTCAGATATCAGGAGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAGTGC 789
Db 361 ATGACTCAGATATCAGGAGGAGTGCCTGGGAGTGCCTGGGAGTGCCTGGGAGTGC 420
Qy 790 GAGGTGAGCTATGGAATCTTCTCCAGAGAGGCTGACAACTTAAAGCAAACTTTGC 849
Db 421 GAGGTGAGCTATGGAATCTTCTCCAGAGAGGCTGACAACTTAAAGCAAACTTTGC 480
Qy 850 AGTAAGCGGAGAGTATGATGACATGCTGACAGATCTGACAGATGAGTATGATC 909
Db 481 AGTAAGCGGAGAGTATGATGACATGCTGACAGATCTGACAGATGAGTATGATC 540
Qy 910 ACTGAGCAAGAGCCCTACACAAAGTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAG 969
Db 541 ACTGAGCAAGAGCCCTACACAAAGTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 595
Qy 970 CATTCCTTT-TATATACGTTTATGGAATGAACTGAAAGAAA 1013

```

Db 596 CAATGCCCTTTATATATATGTTTCTGAATTAAGTGAAGAAA 641

RESULT 15  
AAZ38326  
ID AAZ38326 standard; cDNA; 546 BP.

DT 09-FEB-2000 (first entry)

DE Human transmembrane protein CDNA clone HP10390 coding sequence

KM HP10390; transmembrane domain; stomach cancer cell; antibody;  
 KM assay reagent; diagnostic marker; primer; probe; antisense; gene therapy  
 KM agonist; antagonist; ligand; therapeutic; ds.

OS Homo sapiens.

Key	Location/Qualifiers
FH	
FT	
CDS	1..540

```

FT/      /product= "Human transmembrane protein HP10390"
FT/      /note= "No stop codon given in the specification"
xx

```

PN W09955862-A2

PD 04-NOV-1999.

PF 27-APR-1999; 99WO-JP022226.  
YY

PR 28-APR-1998; 98JP-0119395.  
YY

PA (SAGA ) SAGAMI CHEM RES CENT.  
PA (PROT-) PROTEGENE INC.

PI Kato S, Kimura T;  
yy

DR WPI; 2000-023358/02.  
D-PCDD; 2000-023358/02.

	Human
XX	XX
DE	DE

PT Human proteins with transmembrane domains, involved in control of cell proliferation and differentiation, useful for treating e.g. cancer or inflammation -

PS Claim 3; Page 88; 114pp; English

CC This sequence represents the coding sequence of human CDNA clone  
CC H10390 which encodes a 20 kd protein with one putative transmembrane  
CC domain in the N-terminus. The CDNA was isolated from a human stomach  
CC cancer cell line CDNA library. The protein has no homology with any  
CC known protein. The protein may be used to raise specific antibodies, as  
CC assay reagents, as diagnostic tissue markers, for the isolation of  
CC cognate receptors, ligands and binding proteins, and as biologically  
CC active agents. Nucleotides encoding the protein may be used as primers  
CC and probes or antisense molecules, and in gene therapy. Cells transfected  
CC with these nucleotides may be used to screen for agonists and antagonists  
CC which are potentially useful therapeutically.

**SQ** Sequence 546 BP; 139 A; 126 C; 165 G; 116 T; 0 other;

Query Match	44.2%	Score 472.4	DB 21	Length 546
Best Local Similarity	91.6%	Pred. No. 7.5e-113		
Matches 500	Conservative 0	Mismatches 46	Indels 0	Gaps 0

**Dy**    358 ATGAAGCGCTGGGCGTTCGCTACCCCTACTTTTGGGGGTCCGCTGGGAACATGCCCTGGGCT 417  
**Db**    1 ATGAAGAAGCTGGGCGTTCGCTGCCTCTTGSGGGGCCCTGCTGGAACCGCCTGGGCT 60

Qy 418 CGAAGGACCAAGATCTACACTGTGGAGCTTGCAGGGGCTGTGGATGATTTAGACTGG 477

Db 61 CGAAGGACCAAGATCTCACTGTGGAGCATGCAGGGGCTGTGGATGACATAGAAATCG 120

QY	478	GAATATGCCCCGTGGACCCCAAGAACCAATTCAGATGGGATCCTTCCGAATCAATCCA	537
Db	121	GAATATGCCCCGTGGACCCCAAGAACCAATTCAGATGGGATCCTTCCGAATCAATCCA	180
QY	538	GATGGCAGCCAGTCAGTTGTGGAGTACCTTATGCCGCTCAGAGGCCACCTCAGAGAG	597
Db	181	GATGGCAGCCAGTCAGTTGTGGAGTACCTTATGCCGCTCAGAGGCCACCTCAGAGAG	240
QY	598	TTGCTGTGGAGGTGTGTGACCCGAATAGAGATAGCGGGGAAACAGATGACCTTCTACC	657
Db	241	CTGCTGGAGGAGATATGTGACCCGATAGAGAGATGGGAAACAGATGATCCTTCCACC	300
QY	658	CACCGCAAGAACTACGTACCGGTGTGAGCCGGAAATGAGAAATCCAGTAACCTAGACTTA	717
Db	301	CATCGCAAGAACTACGTACCGGTGTGAGCCGGAAATGAGAAATCCAGTAACCTAGAGACTA	360
QY	718	CAGGCGATCCGAATTTGACTACGATATACGGCGCACCCCTCAAGTTGGCTGTGAGACATT	777
Db	361	CAAGCGATCCCAATTCGACTCGATATTTAGCGGCGACCCCTCAAGTTGGCTGTGAGACATT	420
QY	778	GTGGAGAAATACGAGGATGTGAGCTATGTGAATTCCTTCGAGAGGGGTGACCAAGTTAA	837
Db	421	GTGGAGAAATACGAGGATGTGAGCTATGTGAATTCCTTTCGAGAGGGGTGACCAATGTTAA	480
QY	838	GACAAACTTTTGACATAGCGGACAGATCTATGTGACCATGCCCTGACACAGATCTCAGAT	897
Db	481	GACAAACTTTTGACATAGCGGACAGATCTTTGTGACCATGCCCTGACACATATCGCATGAT	540
QY	898	GAGCTA 903	
Db	541	GAGCTA 546	

Db 121 GAAATTGCCCAGGTGGACCCCAAGAAGACCATTCAGATGGGATCTTCCGGATCAATCCA 180

538 GATGGCAGGCAGTCACTTGTGAGAGTACTTATGCCCCGTCAGAGGGCCCACTCAGAGAG 597  
 |||||||  
 181 GATGGCAGGCAGTCACTGTTGAGAGGTGCTTATGCCCCGTCAGAGGGCCCACTCAGAGAG 240

Db  
241 CTGCTGGAGGAGATATTGTGACCCGGATGAAGGACTATGGGGAAACAGATTGATTCCTTCCACC 300

630 CACCGCAGAGATACGATACGGGTCGTGATGCCCCGAAATCGGAATCAGTGAATAGACTTA / I  
|| |||||||  
Db 301 CATGCCAAGAATACTACGTAGTGTTAGTGGCCCGGAATTGAGAAATCCAGTAGTAATGGAACCTA 360

Db

361 CAAGGCATCCGAATGACCTCAGATAATTAAAGGGCACCCCAAGTTTGGGTGTGAGACATT 420

Db 421 GTGGAGGAATACGAGCATGAACCTCATTTCTTTCCCGAGAGGCTGCACATGTTAAA 480

638 GACAAACCTTGCAGTAAAGCGGACAGATCTAAGACCAATGCCCTGCACAGATCTACAGAT 89  
 |||||  
 481 GACAACTTGCAGTAAAGCGAACAGATCTTGTGACCAATGCCCTGCACATATCGCATGAT 540

W	630	GAGCTA	503
Db	541	GAGCTA	546

Search completed: March 30, 2003, 18:47:39  
Job time : 261.709 secs

Search completed: March 30, 2003, 18:47:39  
Job time : 261.709 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 18:39:57 ; Search time 59.2939 Seconds

(without alignments)  
5529,028 Million cell updates/sec

Title: US-10-082-502-18

Perfect score: 1069  
Sequence: 1 gaattcgccagcaggggggt.....aaaaaatctcgccgcgc 1069

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapept 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	4.3	3581	2	US-08-738-349-1
2	45.6	4.3	7218	1	US-08-232-463-14
3	43.8	4.1	746	4	US-09-013-810-1
4	43.8	4.1	1780	4	US-09-202-5488-5
5	42.4	4.0	3238	4	US-08-123-934A-5
6	42.4	4.0	3238	5	PCT-US94-10080-5
7	42.2	3.9	3410	4	US-09-020-956-110
8	42.2	3.9	3410	4	US-09-030-607-110
9	42.2	3.9	3410	4	US-09-605-785-110
10	42.2	3.9	3410	4	US-09-352-616A-110
11	42.2	3.9	3410	4	US-09-352-616A-110
12	42.2	3.9	3410	4	US-09-352-616A-110
13	42.2	3.9	3410	4	US-09-352-616A-110
14	42.2	3.9	5852	1	US-07-867-106-2
15	42.2	3.9	6200	4	US-09-439-923-1
16	42	3.9	1114	4	US-09-152-060-41
17	42	3.9	1461	5	PCT-US95-04258-4
18	42	3.9	1864	4	US-09-149-476-130
19	42	3.9	289	1	US-08-341-568-3
20	41.6	3.9	289	2	US-08-911-020-3
21	41.6	3.9	1332	2	US-09-057-762-1
22	41.6	3.9	1332	3	US-08-326-119A-1
23	41.6	3.9	976	2	US-08-504-459-9
24	41.4	3.9	2205	3	US-08-888-077A-41
25	41.2	3.8	770	4	US-09-385-982-542
26	41	3.8	1248	2	US-08-897-340-5
27	41	3.8	1248	2	US-08-897-340-5

28	41	3.8	1248	3	US-09-252-329-5	Sequence 5, Appl
29	41	3.8	1315	3	US-09-721-822A-10	Sequence 10, Appl
30	41	3.8	7452	3	US-08-592-500-1	Sequence 1, Appl
31	41	3.8	7452	3	US-08-195-006-1	Sequence 1, Appl
32	41	3.8	7452	5	PCT-US94-07644A-1	Sequence 1, Appl
33	40.8	3.8	458	1	US-08-524-757-1	Sequence 1, Appl
34	40.8	3.8	1502	4	US-09-206-903A-2	Sequence 1, Appl
35	40.8	3.8	1502	4	US-09-206-903A-11	Sequence 11, Appl
36	40.8	3.8	1502	4	US-09-202-122-2	Sequence 2, Appl
37	40.8	3.8	1502	4	US-09-202-122-11	Sequence 11, Appl
38	40.8	3.8	1502	4	US-09-206-935-3	Sequence 2, Appl
39	40.8	3.8	1502	4	US-09-206-935-2	Sequence 3, Appl
40	40.8	3.8	1502	4	US-09-206-936-2	Sequence 2, Appl
41	40.8	3.8	1502	4	US-09-206-936-3	Sequence 3, Appl
42	40.8	3.8	1582	3	US-08-545-196B-10	Sequence 10, Appl
43	40.8	3.8	1582	3	US-08-545-196B-12	Sequence 12, Appl
44	40.6	3.8	253	2	US-08-520-678A-25	Sequence 25, Appl
45	40.6	3.8	253	4	US-08-897-126-25	Sequence 25, Appl

## ALIGNMENTS

RESULT 1  
US-08-738-349-1  
Sequence 1, Application US/08738349  
Patent No. 5869638  
GENERAL INFORMATION:  
APPLICANT: Takashita, Sunao  
APPLICANT: Okazaki, Makoto  
APPLICANT: Kawai, Shinji  
APPLICANT: Tsujimura, Atsushi  
APPLICANT: Amann, Egon  
TITLE OF INVENTION: Bone-Related Cadherin-Like Protein and  
TITLE OF INVENTION: Process for Its Production  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dinner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/738,349  
FILING DATE: 25-OCT-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/364,439  
FILING DATE:  
APPLICATION NUMBER: US 08/112,061  
FILING DATE: 26-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Barker, M. P.  
REGISTRATION NUMBER: 32,013  
REFERENCE/DOCKET NUMBER: 02481.1323-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4000  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3581 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:

ORGANISM: Mus musculus  
STRAIN: osteoblastic cell line MC3T3E1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 284..2671  
US-08-738-349-1

Query Match 4.3%; Score 46; DB 2; Length 3581;  
Best Local Similarity 63.6%; Pred. No. 0.0031;  
Matches 70; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Y 960 AAGATGGCAGCATGCTTTATATAGCTTTTATGGAATGAACTGAAAACTCTT 1019  
Db 3466 AAGAAGCACTTTGAAATTAATAAAGATCTTTTAAAAAATAAAAAA 3525

Y 1020 GAACCCAGATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1069  
Db 3526 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 3575

RESULT 2  
US-08-232-463-14/c  
Sequence 14, Application US/08232463  
Patent No. 5670367

GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEFFELINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOXLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232.463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935.313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)836-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZ9pt-f18  
US-08-232-463-14

Query Match 4.3%; Score 45.6; DB 1; Length 7218;  
Best Local Similarity 5.0%; Pred. No. 0.0057;  
Matches 21; Conservative 219; Mismatches 178; Indels 0; Gaps 0;

Y 643 ATTGACCTTCTACCCAGCAAGTACGATGCGCTGAGCCGGAAGTGAATCC 702  
Db 1481 AATTACCTATCTATCACTAGTAAAGATGAAGAAATTTGTAACRRRRRRRR 1422  
Y 703 AGTGAACTAGACTTACGGGCAATCGAATGACTGATGATGAGCGGCCCTCAAGTT 762  
Db 1421 RRR 1362  
Y 763 GCGTGTGAGAGCATGTTGGAAGATGAGATGAGCTTATGCAATTCCTCCAGAG 822  
Db 1361 RRR 1302  
Y 823 GCTGACAGCTTAAGACAACTTGCTAGTAAAGGACAGATCTATGACCATGCCCTG 882  
Db 1301 RRR 1242  
Y 883 CACAGATCTCAGATGAGCTATGATCACTGAGCAGACGAGCTCACCAAGTGAATG 942  
Db 1241 RRR 1182  
Y 943 AACACCCCGAGGAGGAGATGCGCATGCTTTATATAGCTTTTATGGAATG 1002  
Db 1181 RRR 1122  
Y 1003 AACTGAAAAAACTTTGAAACCGAAGTAAAAAATAAATAAATAAATAAATTC 1060  
Db 1121 RRR 1064

RESULT 3  
US-09-013-810-1  
Sequence 1, Application US/09013810  
Patent No. 6197551

GENERAL INFORMATION:  
APPLICANT: Busfield, Samantha J.  
TITLE OF INVENTION: No. 6197551e1 TANGO 80 Molecules and Uses Thereof  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/013.810  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragoras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MEI-010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 746 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 135..428







Db 3236 TTATTAGCGGGGTGAATTTATCTAAGTGACATGAGTATATGTTATG 3295  
QY 964 TGGCAGCATTCCTTTTATATTAAGCTTTTATGAAATGAACTGAA 1023  
Db 3296 TGACAAATTTAAAGCTTCTATATGTTTAAAAA 3355  
QY 1024 CCGAAAGTAAAAA 1056  
Db 3356 AAAAAAAAAAAAAAAAAAAAAA 3388

## RESULT 9

US-09-605-785-110  
Sequence 110, Application US/09605785  
Patent No. 6321716  
GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darlick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aljun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C16  
CURRENT APPLICATION NUMBER: US/09/605,785  
CURRENT FILING DATE: 2000-06-27  
NUMBER OF SEQ ID NOS: 835  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 110  
LENGTH: 3410  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-605-785-110

Query Match 3.9%; Score 42.2; DB 4; Length 3410;  
Best Local Similarity 54.2%; Pred. No. 0.036;

Matches 83; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

QY 904 TGAATCACTGAGCAAGCAGCTACACCAACGATGAGAACCCCGAGAGGAGAGA 963  
Db 3236 TTATTAGCGGGGTGAATTTATCTAAGTGACATGAGTATATGTTATG 3295  
QY 964 TGGCAGCATTCCTTTTATATTAAGCTTTTATGAAATGAACTGAA 1023  
Db 3296 TGACAAATTTAAAGCTTCTATATGTTTAAAAA 3355  
QY 1024 CCGAAAGTAAAAA 1056  
Db 3356 AAAAAAAAAAAAAAAAAAAAAA 3388

## RESULT 10

US-09-439-313-110  
Sequence 110, Application US/09439313  
Patent No. 6329505  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan Louise

APPLICANT: Jiang Yuqi  
APPLICANT: Reed, Steven G.  
APPLICANT: Kalos, Michael  
APPLICANT: Fanger, Gary  
APPLICANT: Retter, Mark  
APPLICANT: Stolk, John  
APPLICANT: Day, Craig  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C9  
CURRENT APPLICATION NUMBER: US/09/439,313  
CURRENT FILING DATE: 1999-11-12  
NUMBER OF SEQ ID NOS: 575  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 110  
LENGTH: 3410  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-439-313-110

Query Match 3.9%; Score 42.2; DB 4; Length 3410;  
Best Local Similarity 54.2%; Pred. No. 0.036;

Matches 83; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

QY 904 TGAATCACTGAGCAAGCAGCTACACCAACGATGAGAACCCCGAGAGGAGAGA 963  
Db 3236 TTATTAGCGGGGTGAATTTATCTAAGTGACATGAGTATATGTTATG 3295  
QY 964 TGGCAGCATTCCTTTTATATTAAGCTTTTATGAAATGAACTGAA 1023  
Db 3296 TGACAAATTTAAAGCTTCTATATGTTTAAAAA 3355  
QY 1024 CCGAAAGTAAAAA 1056  
Db 3356 AAAAAAAAAAAAAAAAAAAAAA 3388

## RESULT 11

US-09-352-616A-110  
Sequence 110, Application US/09352616A  
Patent No. 6395278  
GENERAL INFORMATION:

APPLICANT: Dillon, Davin C.  
APPLICANT: Harlocker, Susan Louise  
APPLICANT: Jiang, Yuqi  
APPLICANT: Xu, Jiangchun  
APPLICANT: Mitcham, Jennifer Lynn  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE  
FILE REFERENCE: 210121.427C8  
CURRENT APPLICATION NUMBER: US/09/352,616A  
CURRENT FILING DATE: 1999-07-13  
NUMBER OF SEQ ID NOS: 472  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 110  
LENGTH: 3410  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-352-616A-110

Query Match 3.9%; Score 42.2; DB 4; Length 3410;  
Best Local Similarity 54.2%; Pred. No. 0.036;

Matches 83; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

QY 904 TGAATCACTGAGCAAGCAGCTACACCAACGATGAGAACCCCGAGAGGAGAGA 963  
Db 3236 TTATTAGCGGGGTGAATTTATCTAAGTGACATGAGTATATGTTATG 3295  
QY 964 TGGCAGCATTCCTTTTATATTAAGCTTTTATGAAATGAACTGAA 1023  
Db 3296 TGACAAATTTAAAGCTTCTATATGTTTAAAAA 3355  
QY 1024 CCGAAAGTAAAAA 1056

DB 3356 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3388

## RESULT 12

US-09-602-877A-100  
Sequence 100, Application US/09602877A  
Patent No. 6432707  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.446CS  
CURRENT APPLICATION NUMBER: US/09/602.877A  
CURRENT FILING DATE: 2000-06-22  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 100  
LENGTH: 3410  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-602-877A-100

Query Match 3.9%; Score 42.2; DB 4; Length 3410;  
Best Local Similarity 54.2%; Pred. No. 0.036;  
Matches 83; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

QY 904 TGAATCACTGGAGCAGACGCTTACACCAACGATGGAACCCCGAGGGGGAAGA 963  
DB 3236 TTATTACGGGGGTGAATTTTATCTGTAGTGAATGAGTATGTTATG 3295  
QY 964 TGCAGCATTCCTTTATTTATGCTTTTATGAAATGAACTGAAACCTCTGAAA 1023  
DB 3296 TGACAAATTAAGGCTTCTTATGTTTAAAAA 3355  
QY 1024 CCGAAAGTAAAAA 1056  
DB 3356 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3388

## RESULT 13

US-09-232-149A-110  
Sequence 110, Application US/09232149A  
Patent No. 6465611  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer Lynn  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE  
FILE REFERENCE: 210121.427C6  
CURRENT APPLICATION NUMBER: US/09/232.149A  
CURRENT FILING DATE: 1999-01-15  
NUMBER OF SEQ ID NOS: 338  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 110  
LENGTH: 3410  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-232-149A-110

Query Match 3.9%; Score 42.2; DB 4; Length 3410;  
Best Local Similarity 54.2%; Pred. No. 0.036;  
Matches 83; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

QY 904 TGAATCACTGGAGCAGACGCTTACACCAACGATGGAACCCCGAGGGGGAAGA 963  
DB 3236 TTATTACGGGGGTGAATTTTATCTGTAGTGAATGAGTATGTTATG 3295  
QY 964 TGCAGCATTCCTTTATTTATGCTTTTATGAAATGAACTGAAACCTCTGAAA 1023

DB 3296 TGACAAATTAAGGCTTCTTATGTTTAAAAA 3355  
QY 1024 CCGAAAGTAAAAA 1056  
DB 3356 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3388

## RESULT 14

US-07-867-106-2/C  
Sequence 2, Application US/07867106  
Patent No. 5389526  
GENERAL INFORMATION:  
APPLICANT: Slade, Martin B  
APPLICANT: Chang, Andy C M  
APPLICANT: Williams, Keith L  
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular  
TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526rls  
STREET: One Liberty Place 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/867.106  
FILING DATE: 19920625  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PJ 7187  
APPLICATION NUMBER: PCT/AU90/00530  
FILING DATE: 02-NOV-1989

ATTORNEY/AGENT INFORMATION:  
NAME: Feeney, Joanne Longo  
REGISTRATION NUMBER: 35,134  
REFERENCE/DOCKET NUMBER: RICE-0002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 5852 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ANTI-SENSE: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 2378..5038

NAME/KEY: CDS

LOCATION: 2378..5038

US-07-867-106-2

Query Match 3.9%; Score 42.2; DB 1; Length 5852;  
Best Local Similarity 70.9%; Pred. No. 0.047;  
Matches 56; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 978 TTATATTACGTTTATGGAATGAACTCTTGAACGAAAGTAAAAA 1037  
DB 5682 TTATATTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 5623  
QY 1038 AAAAAAAAAAAAAAAAAA 1056  
DB 5622 AAAAAAAAAAAAAAAAAA 5604

```

RESULT 15
US-09-439-923-1
; Sequence 1, Application US/09439923
; Patent No. 6426208
; GENERAL INFORMATION:
; APPLICANT: Emily D. Kakkis
; APPLICANT: Becky Tanamachi
; TITLE OF INVENTION: Recombinant Alpha-L-Iduronidase, Methods
; TITLE OF INVENTION: for Producing and Purifying the Same and Methods for
; FILE REFERENCE: 08000510500
; CURRENT APPLICATION NUMBER: US/09/439,923
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 6200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1558)...(3516)
US-09-439-923-1

```

```

Query Match          3.9%; Score 42.2; DB 4; Length 6200;
Best Local Similarity 67.8%; Pred. No. 0.049;
Matches 59; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY  970 CATGCCCTTTATATTACGTTTATGGAATGAACTGAAACCTTGAAACCGAAA 1029
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  3608 CTTGCAATATATTTTATATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  1030 GTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  3668 AAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: March 30, 2003, 20:59:50  
 Job time : 104.294 secs





[illegible]

PRIOR FILING DATE:	1998-03-12
PRIOR APPLICATION NUMBER:	60/078911
PRIOR FILING DATE:	1998-03-20
PRIOR APPLICATION NUMBER:	60/079294
PRIOR FILING DATE:	1998-03-25
PRIOR APPLICATION NUMBER:	60/079665
PRIOR FILING DATE:	1998-02-27
PRIOR APPLICATION NUMBER:	60/079728
PRIOR FILING DATE:	1998-03-27
PRIOR APPLICATION NUMBER:	60/080165
PRIOR FILING DATE:	1998-03-31
PRIOR APPLICATION NUMBER:	60/081200
PRIOR FILING DATE:	1998-04-09
PRIOR APPLICATION NUMBER:	60/081222
PRIOR FILING DATE:	1998-04-09
PRIOR APPLICATION NUMBER:	60/081695
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/081811
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/081815
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/082939
PRIOR FILING DATE:	1998-04-24
PRIOR APPLICATION NUMBER:	60/083322
PRIOR FILING DATE:	1998-04-28
PRIOR APPLICATION NUMBER:	60/083545
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/084600
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/084622
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/084633
PRIOR FILING DATE:	1998-05-13
PRIOR APPLICATION NUMBER:	60/085145
PRIOR FILING DATE:	1998-05-12
PRIOR APPLICATION NUMBER:	60/085322
PRIOR FILING DATE:	1998-05-13
PRIOR APPLICATION NUMBER:	60/085333
PRIOR FILING DATE:	1998-05-13
PRIOR APPLICATION NUMBER:	60/085575
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085697
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085700
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/086414
PRIOR FILING DATE:	1998-05-22
PRIOR APPLICATION NUMBER:	60/086430
PRIOR FILING DATE:	1998-05-22
PRIOR APPLICATION NUMBER:	60/087100
PRIOR FILING DATE:	1998-05-28
PRIOR APPLICATION NUMBER:	60/088022
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/088730
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/088855
PRIOR FILING DATE:	19/98-06-11
PRIOR APPLICATION NUMBER:	60/089533
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089599
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089907
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089944
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/090349
PRIOR FILING DATE:	1998-06-23



PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090538  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07

Query Match 64.6%; Score 691; DB 9; Length 1210;  
Best Local Similarity 84.4%; Pred. No. 1,7e-187;  
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

QY 12 GAGGGGGGCTCTGCTGCTGCGAGGGGCTCTAAAGCTGCTGCGCGAGAGTTTG 71  
DB 196 GACCGGGGCTCTGCTGCTGCGAGGGGCTCTAAAGCTGCTGCGCGAGAGTTTG 254  
QY 72 GAGGGGGGCTCTGCTGCTGCGAGGGGCTCTAAAGCTGCTGCGCGAGAGTTTG 129  
DB 255 GAGGGGGGCTCTGCTGCTGCGAGGGGCTCTAAAGCTGCTGCGCGAGAGTTTG 314  
QY 130 GCTTGGTGGCTTAAAGTGGGGAAGCTGGGCGAGCGGTC-ACGCGCACCTGTAAC 188  
DB 315 GCTTGGTGGCTTAAAGTGGGGAAGCTGGGCGAGCGGTC-ACGCGCACCTGTAAC 374  
QY 189 CTGGCGGAGCGGCGGCTCTGGAAGCGGCGGAGCGGCGGCTCTGGAAGCGGCGG 248  
DB 375 GTGGCAGGCGGCGGAGCGGCTCTGGAAGCGGCGGAGCGGCGGCTCTGGAAGCGG 434  
QY 249 CTAGGGTGGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAGCGG 308  
DB 435 -TAGTTGGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAGCGG 493  
QY 309 ACGGGGGGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 368  
DB 494 ACGGGGGGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 553  
QY 369 GGGTGGCTAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 428  
DB 554 GGGTGGCTAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 613  
QY 429 AGATCTAGCTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 488  
DB 614 AGATCTAGCTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 673  
QY 489 CGTGGAGCGGCGGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 548  
DB 674 GGTGGAGCGGCGGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 733  
QY 549 GTCAGTGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 608  
DB 734 GTCAGTGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 793  
QY 609 GGTGTGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 668  
DB 794 GATATGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 853  
QY 669 CTAGTGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 788  
DB 854 CTAGTGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 913  
QY 729 AATGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 788  
DB 914 AATGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 973  
QY 789 CGAGTGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 848

DB 974 CGAGTGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 1033  
QY 849 CAGTGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 908  
DB 1034 CAGTGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 1093  
QY 909 CAGTGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 968  
DB 1094 CAGTGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 1148  
QY 969 GCATGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 1013  
DB 1149 GCATGTGAGCGGCTCTGGAAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAG 1195

## RESULT 2

US-10-121-049-173

Sequence 173, Application US/10121049  
Publication No. US2003002239A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Geriltsen, Mary E.  
APPLICANT: Goddard, Audrey E.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zhen  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C17  
CURRENT APPLICATION NUMBER: US/10/121, 049  
CURRENT FILING DATE: 2002-04-12  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-121-049-173

Query Match 64.6%; Score 691; DB 9; Length 1210;  
Best Local Similarity 84.4%; Pred. No. 1,7e-187;  
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

QY 12 GAGGGGGGCTCTGCTGCTGCGAGGGGCTCTAAAGCTGCTGCGCGAGAGTTTG 71  
DB 196 GACCGGGGCTCTGCTGCTGCGAGGGGCTCTAAAGCTGCTGCGCGAGAGTTTG 254  
QY 72 GAGGGGGGCTCTGCTGCTGCGAGGGGCTCTAAAGCTGCTGCGCGAGAGTTTG 129  
DB 255 GAGGGGGGCTCTGCTGCTGCGAGGGGCTCTAAAGCTGCTGCGCGAGAGTTTG 314  
QY 130 GCTTGGTGGCTTAAAGTGGGGAAGCTGGGCGAGCGGTC-ACGCGCACCTGTAAC 188  
DB 315 GCTTGGTGGCTTAAAGTGGGGAAGCTGGGCGAGCGGTC-ACGCGCACCTGTAAC 374  
QY 189 CTGGCGGAGCGGCGGCTCTGGAAGCGGCGGAGCGGCGGCTCTGGAAGCGGCGG 248  
DB 375 GTGGCAGGCGGCGGAGCGGCTCTGGAAGCGGCGGAGCGGCGGCTCTGGAAGCGG 434  
QY 249 CTAGGGTGGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAGCGG 308  
DB 435 -TAGTTGGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTCTGGAAGCGG 493

OY	309	ACGGGGGGGGCGGAGGAGCAGAGGTGGAGGACCGCTGTACACATAAGATGAAAGGCTG	368
OY	309	ACGGGGGGGGCGGAGGAGCAGAGGTGGAGGACCGCTGTACACATAAGATGAAAGGCTG	368
Db	484	ACGGGGGGGGCGCTGGAGGAGCAGAGGTGGAGGACCGCTGTACACATAAGATGAAAGGCTG	553
OY	369	GGGTGGGCTTACCCCTACTTCTTGGGGGGCTCTGCTGGGAAGTGGCTGGGCTGGAAGGACA	428
Db	554	GGGTGGGCTTACCCCTACTTCTTGGGGGGCGCTGGCTGGGAAGCGGCTGGGCTGGAGGACCA	613
OY	429	AGATCTACACGTGGAGCTTTCAGAGGCTCTGGTGGATGAAATTAGAGTGGGAAATTGGCCG	488
Db	614	GGATCTCCACGTGGAGAGATGCAAGGCTCTGGTGGATGAAATGAAATGGGAAATTGGCCA	673
OY	489	CGTGGACCCCAAGAAAGACATTCAGATGGGATCCCTCCGATCAATTCAGATGGACGCA	548
Db	674	GGTGGACCCCAAGAAAGACATTCAGATGGGATCTTCTCCGATCAATTCAGATGGACGCA	733
OY	549	GTCAGTTGTGAGAGGTACCTTATGCCCCGCTCAGAGGCCACCTCAGACAGATTGCTTGAGA	608
Db	734	GTCAGTTGTGAGAGGTACCTTATGCCCCGCTCAGAGGCCACCTCAGACAGAGCTGCTGGAGA	793
OY	609	GGTGTGTACCCGAAATGAAAGAGTACGGGGGAAACAGATTGACCTTCTTACCCACCGCAAGAA	668
Db	794	GATATGTACCGGAGTGAAGAGATGAGGGGAAACAGATTGATCTTCCACCCATGCGCAAGAA	853
OY	669	CTACGTACGCTCTGTGAGCCCGGATGGAGAATCCAGTGAAGTACTTACAGGSCATCCG	728
Db	854	CTACGTACGCTTACTGGGGCCCGAATGGAGAATCCAGTGAAGTACTTACAGGSCATCCG	913
OY	729	AATTGACTCAGATATCAGCGGACCCCTCAAGTTGCGTGTGAGAGCATTTGTGGAGAATA	788
Db	914	AATCAGCTCAGATATTTAGCGGACCCCTCAAGTTGCGTGTGAGAGCATTTGTGGAGAATA	973
OY	789	CGAGGATAGGTTTTCGAATCTCTCCAGAGAGCGTGACAAGCTTAACACAACTTGG	848
Db	974	CGAGGATAGGTTTTCGAATCTTCTCCGAGAGCGTGACAAGCTTAACACAACTTGG	1033
OY	849	CAGTAAAGCGGACAGATCTATGTGACATGCGCTGACAGACATCTCAGCATGAGCTATGAAT	908
Db	1034	CAGTAAAGCGGACAGATCTTGTGACATGCGCTGACAGATCTCAGCATGAGCTATGAAC	109
OY	909	CACGTGAGGACAAAGACCTTACACCAACGATGGAACACCCCGAGAGGGGAAGATGGCA	968
Db	1094	CACGTGAGGACACCCAC-----ACTGGCGTTGATGGATCACCCCGAGGGGGAAATGGTG	114
OY	969	GCATTGCGCTTT--TATATTACGTTTTTATGGAATGAACGAAAAAA	1013
Db	1149	GCATTGCGCTTTATATTATTATGTTTTTACTGAAATTAACGAAAAA	1195
RESULT 3			
US-10-123-904-173			
: Sequence 173, Application us/101123904			
: Publication No. US20030022328A1			
: GENERAL INFORMATION:			
: APPLICANT: Baker, Kevin P.			
: APPLICANT: Beresini, Maureen			
: APPLICANT: DeForge, Laura			
: APPLICANT: Desnoyers, Luc			
: APPLICANT: Filvaroff, Ellen			
: APPLICANT: Gao, Wei-Qiang			
: APPLICANT: Gerritsen, Mary E.			
: APPLICANT: Goddard, Audrey			
: APPLICANT: Godowski, Paul J.			
: APPLICANT: Gurney, Austin L.			
: APPLICANT: Sherwood, Steven			
: APPLICANT: Smith, Victoria			
: APPLICANT: Stewart, Timothy A.			
: APPLICANT: Tumas, Daniel			
: APPLICANT: Watanabe, Collin K			
: APPLICANT: Wood, William			
: APPLICANT: Zhang, Zemin			
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			

	Query Match	Best Local Match	Similarity	84.4%	Score 691	DB 9	Length 1210	
	Matches 850	Conservative	0	Mismatches 145	Indels 12	Gaps		
Qy	12	GAGGGGGCTCTCCGCTCGCTCGAGGCGCGCTCTTAAAGCTGCGTCCGCGGAGAGTTTG	71	11	111111	1111	1111	1111
Db	196	GACCGGGGCTCTTGGCTGTGCGTGGAGGCTCTCT-AGCTGCTGTTCGGCGAGAGTTTG	254	11	111111	1111	1111	1111
Qy	72	GAGGGGGGGCTTAAAGGTCAGTTTCGGTGGGGGCTCGACGG--GACCTCAGATCTCC	129	11	111111	1111	1111	1111
Db	255	GAGGGGGGGCTTAAAGGTCAGTTTCGGTGGGGGCTCGACGG--GACCTCAGATCTCC	314	11	111111	1111	1111	1111
Qy	130	GCTTAAAGTCTCAGTTAAAGTGGCGGAGAGCTGGGCGAGGGGCTC-AGTGGCCACCTCGAAG	188	11	111111	1111	1111	1111
Db	315	GCTTAAAGTACAGTTAAAGTGGCGGAGAGCTGGGCGAGGGGCTC-AGTGGCCACCTCGAAG	374	11	111111	1111	1111	1111
Qy	189	CTGGCGGAGCGGAGCGCTCTGAGAAACCCGGGACAGCCGCTTTTCCAGCCAGCTG	248	1111	1111	1111	1111	1111
Db	375	GTGGCAGCGGGGAGAGCGCTCTGAGAAACCCGGGACAGCCGCTTTTCCAGCCAGCTG	434	1111	1111	1111	1111	1111
Qy	249	CTAAGGTTGGGACCCACAGAAAAACAAGTGAAGTCCGGCTCTTCCAGAGCTGGGCC	308	1111	1111	1111	1111	1111
Db	435	-TAGTTTCGGGAGCCACTGCGGCCAAAGTGAGTCCAGCGGCTTCCAGCGCTGGGCC	493	1111	1111	1111	1111	1111
Qy	309	ACGGGGGGGGCGCGCTGGGAGACAGAGTGGAGCGACCTCTTACACTTAAATGAAGCTG	368	1111	1111	1111	1111	1111
Db	494	ACGGGGGGGGCGCGCTGGGAGACAGAGTGGAGCGACCTCTTACACTTAAATGAAGCTG	553	1111	1111	1111	1111	1111
Qy	369	GGTGGGCTAGACCTACTTTTGGGGGCTCTGCTGGGAACTGCTGGGGCTCGAAGGAGCA	428	1111	1111	1111	1111	1111
Db	554	GGTGGGCTAGACCTACTTTTGGGGGCTCTGCTGGGAACTGCTGGGGCTCGAAGGAGCA	613	1111	1111	1111	1111	1111
Qy	429	AGATCTACACTGTGAGCTTTCGAGGGCTCTGCTGGAATGAATTAAGTGGGAAATTCGCCG	488	1111	1111	1111	1111	1111
Db	614	GGATCTACACTGTGAGCTTTCGAGGGCTCTGCTGGAATGAATTAAGTGGGAAATTCGCCG	673	1111	1111	1111	1111	1111
Qy	489	CGTGGACCCCAAGAAAGACATTCAGATGGGATCTTCCGATCAATCCAGATGGACGCA	548	1111	1111	1111	1111	1111
Db	674	GGTGGACCCCAAGAAAGACATTCAGATGGGATCTTCCGATCAATCCAGATGGACGCA	733	1111	1111	1111	1111	1111
Qy	549	GTACAGTTGGAGAGTACCTTATGCCGGCTCAGAGGGGCCACTCACAGATTCGTTGAGGA	608	1111	1111	1111	1111	1111
Db	734	GTACAGTTGGAGAGTACCTTATGCCGGCTCAGAGGGGCCACTCACAGATTCGTTGAGGA	793	1111	1111	1111	1111	1111
Qy	609	GGTGTGTGACCAATGAAGAGTACGGGGACAGATTGAACCTTTCACCCACCGCAAGAA	668	1111	1111	1111	1111	1111
Db	794	GATATGTGACCGGATGAAGAGTATGGGGACAGATTCCTTCACCCACCGCAAGAA	853	1111	1111	1111	1111	1111
Qy	669	CTACGTAGCGCTCTGAGACCGGAAATGAGAAATCCAGTGAATTAAGTTCAGAGGCATCCG	728	1111	1111	1111	1111	1111
Db	854	CTACGTAGCGCTCTGAGACCGGAAATGAGAAATCCAGTGAATTAAGTTCAGAGGCATCCG	913	1111	1111	1111	1111	1111
Qy	729	AATTGACACAGATTCACAGCGGACACCTCAAGTTTCGCTGTGAGACATTCGTGGAGAAAT	788	1111	1111	1111	1111	1111
Db	914	AATTGACACAGATTCACAGCGGACACCTCAAGTTTCGCTGTGAGACATTCGTGGAGAAAT	973	1111	1111	1111	1111	1111
Qy	789	CGAGATATGAGCTTATGAATTTCTTCCAGAGAGGCTGCAACGTTTAAAGCAAACTTTG	848	1111	1111	1111	1111	1111
Db	974	CGAGATATGAGCTTATGAATTTCTTCCAGAGAGGCTGCAACGTTTAAAGCAAACTTTG	1033	1111	1111	1111	1111	1111
Qy	849	CAGTAAGGGAGATCTATGTGACATGCGCTCGACAGATTCAGATGAGCTATGAT	908	1111	1111	1111	1111	1111

DB 1034 CAGTAAAGGACAGATCTTTTGTGACCATGCCCCGACATATGCGATGTGAGCTATGAC 1093  
QY 909 CACTGGAGCAAGCAGCCCTACACCAACCTGATGGAACACCCCGAGAGGGAAGATGCA 968  
DB 1094 CACTGAGACAGCCAC-----ACTGGCTGTATGATGATCAGCCCGAGAGGGAATGTG 1148  
QY 969 GCATTGCTTT--TATATTAGCTTTTATGGAATGAATGATAAAAA 1013  
DB 1149 GCAATGCTTTTATATATATATATATGTTTACTGMAATTAAGTAAAAA 1195

## RESULT 4

US-10-140-470-173  
; Sequence 173, Application US/10140470  
; Publication No. US2003002231A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P330R1C160  
; CURRENT APPLICATION NUMBER: US/10/140,470  
; CURRENT FILING DATE: 2002-05-06  
; PRIOR APPLICATION removed - See Palm or File Wrapper  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-140-470-173

## Query Match

Best Local Similarity 84.4%; Score 691; DB 9; Length 1210;  
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

QY 12 GAGGGGGGTCCTGCTGCTCGGAGGCGCTTAAGCTGCTGCTGCGGCGAGATTG 71  
DB 196 GACCGGGGTCCTGCTGCTGCTGCGAGCCTCTCT-AGCTGCTGCTGCGCGAGATTG 254  
QY 72 GAGGGGGGCTTAGGGCTAGTTGCTGCTGCGGCGCTGCGACG--GACCTCAGATCTCC 129  
DB 255 GAGGGGGGCTTAGGGCTAGTTGCTGCTGCGGCGCTGCGACGCGACGAGCTGAGTCCC 314  
QY 130 GCTTAGTCCCTAGTAAGTGGCGGGAAGCTGCGGCGGCTG-AGTGGCCACCTGTAAC 188  
DB 315 GCTTAGTCCCTAGTAAGTGGCGGGAAGCTGCGGCGGCTGCGGCGGACACCCGCTGT 374  
QY 189 GTGGGGGAGCGGAGCGCTGAGAAACCGGAGACGCGCGCTTTTCCGACGCACTG 248  
DB 375 GTGGCGAGGGGAGCGCTGAGAAACCGGAGACGCGCTGCTGCGACGCGAGTGG 434  
QY 249 CTAAGGTTGGGACCCACAGAAAAGTGAAGTCCGCTGCTTCCAGAGCTGGGCC 308  
DB 435 -TAGTTTGGGAGCGACTGGGGCCAAAGTGAAGTCAACGCGCTTCCAGCGCTGGGCC 493  
QY 309 ACGCGGGGCGCTGGAGACAGAGTGAAGTCAACCTGTTAAGTAAGTAAGGCTG 368  
|||||

DB 494 ACGCGGGGCGCTGGAGACAGAGTGGAGGAGCCCATTTACGCTMAAGTAAGGCTG 553  
QY 369 GGGTTGGCTACCTCTACTTTTGGGGCTCTCTGGGAATGCTGGGCTCGAAGAGCCA 428  
DB 554 GGGTTGGCTGGCGCTCTCTCTCTCTGGGGCGCTGCTGGGAACCGCTGGGCTCGAAGAGCCA 613  
QY 429 AGATCTACACTGTGAGCTTGCAGGGCTCTGGTGGATGAATTAAGTGGGAATGCGCCG 488  
DB 614 GGATCTCAGCTGTGAGCATTCAGAGGCTCTGTGTGATGAATTAAGTGGGAATGCGCCA 673  
QY 489 CGTGACCCCAAGAACCATTCAGATGGATCCCTCCGATCAATCAGATGCGACCCA 548  
DB 674 GGTGAGCCCAAGAACCATTCAGATGGATCCCTCCGATCAATCAGATGCGACCCA 733  
QY 549 GTGAGTTGGAGGTGAGCTTATGCCCCCTCAGAGAGCCCACTCAGAGTTCCTTGAGA 608  
DB 734 GTGAGTTGGAGGTGAGCTTATGCCCCCTCAGAGAGCCCACTCAGAGTTCCTTGAGA 793  
QY 609 GGTGTGTAGCCGATGAGAGAGTACGGGGAACAGATGACCTTCTACCGACGCAAGAA 668  
DB 794 GATATGTAGCCGATGAGAGAGTATGGGGAACAGATGATTCCTCCACCATGCAAGAA 853  
QY 669 CTAGTACGCGCTGTGAGCCGGAATGAGAAATCCAGTGAATCTACAGGGCATCCG 728  
DB 854 CTAGTACGCTGTGAGCCGGAATGAGAAATCCAGTGAATCTACAGGGCATCCG 913  
QY 729 AATTGACTGATATGAGGCGGCAACCTCAGATTGGCTGAGAGCATTTGGAAGATA 788  
DB 914 AATGACTGATATGAGGCGGCAACCTCAGATTGGCTGAGAGCATTTGGAAGATA 973  
QY 789 CGAGGATGACTTATGAAATCTCTCCAGAGAGGCTGACAGATTAAGACAACTTTG 848  
DB 974 CGAGGATGACTTATGAAATCTCTCTCCAGAGAGGCTGACAGATTAAGACAACTTTG 1033  
QY 849 CAGTAAGCGGAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 908  
DB 1034 CAGTAAGCGGAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1093  
QY 909 CACTGAGCAAGCAGCCCTACCAACGATGATGATGATGATGATGATGATGATGATGAT 968  
DB 1094 CACTGAGCAAGCAGCCCTACCAACGATGATGATGATGATGATGATGATGATGATGAT 1148  
QY 969 GCATTGCTTT--TATATTAGCTTTTATGGAATGAATGATAAAAA 1013  
DB 1149 GCAATGCTTTTATATATATATATATGTTTACTGMAATTAAGTAAAAA 1195  
|||||

## RESULT 5

US-10-175-746-173  
; Sequence 173, Application US/10175746  
; Publication No. US2003002270A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P330R1C353  
; CURRENT APPLICATION NUMBER: US/10/175,746



```
Db 554 GGGTGGCTGGCCCTGCTTCTGGGGGGCCCTGCTGGGAACCGGCTGGGCTGGAGGAGCCA 613
Qy 429 AGATCTACACTGTGTGAGCTTTCAGAGGCTCTGTGGATGAATTAGAGGGAATTTGCCCG 488
Db 614 GGAATCTCACTGTGTGAGCAATGAGAGGCTGTGTGTGATGAATGAATGGGAATTTGCCCA 673
Qy 489 CGTGGACCCCAAGAAAGCAATTCAGATGGATCTTCCGAATCAATCAATGGCAGCCA 548
Db 674 GGTGGACCCCAAGAAAGCAATTCAGATGGATCTTCCGAATCAATCAATGGCAGCCA 733
Qy 549 GTCAGTGTGTGAGTACCTTATGCCCCCTCAGAGAGCCACCTCAGAGATGTGTGGTGA 608
Db 734 GTCAGTGTGTGAGTGTGCTTATGCCCCCTCAGAGAGCCACCTCAGAGATGTGTGGTGA 793
Qy 609 GGTGTGTGACCGAATGAGAGATGAGGGGAACAGATTGACCTTTCACCCAGCCAGAA 668
Db 794 GATATGTGACCGAATGAGAGATGAGGGGAACAGATTGATCTTCCAGCCATGGCAAGAA 853
Qy 669 CTAGTACGGGTGTGTGAGCCCGGAATGGAATCCAGTGAATGACTTACAGGGCATCCG 728
Db 854 CTAGTACGGGTGTGTGAGCCCGGAATGGAATCCAGTGAATGACTTACAGGGCATCCG 913
Qy 729 AATTGACTCAGATATCAGCGCACCCCTCAATGTTGGGTGAGAGGATGTGGAGATA 788
Db 914 AATGCACTCAGATATTAAGCGCACCCCTCAATGTTGGGTGAGAGGATGTGGAGATA 973
Qy 789 CGAGATGAGCTTATGCAATTTCTTCCAGAGAGGCTGACACAGTTAAAGACAACCTTG 848
Db 974 CGAGATGAGCTTATGCAATTTCTTCCAGAGAGGCTGACACAGTTAAAGACAACCTTG 1033
Qy 849 CAGTAAGCGGAGATCTATGTGACATGCCCGCACAGATCTCAGATGATGATGAAT 908
Db 1034 CAGTAAGCGGAGATCTATGTGACATGCCCGCACAGATCTCAGATGATGATGAAT 1093
Qy 909 CACTGAGCAAGCAGCTTACACCAACCTGTGATGAACACCCAGAGAGGAGATGACA 968
Db 1094 CACTGAGCAAGCAGCTTACACCAACCTGTGATGAACACCCAGAGAGGAGATGACA 1148
Qy 969 GCATTCGCCCTT--TATTTACGTTTTTATGGAATGACTGAAAAA 1013
Db 1149 GCATTCGCCCTTATATATATATGTTTTTACTGAAATTAAGTAAAAA 1195

RESULT 7
; Sequence 173, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C288
; CURRENT APPLICATION NUMBER: US/10-0176,921
; CURRENT FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
```

```
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-921-173

Query Match      64.6%; Score 691; DB 9; Length 1210;
Best Local Similarity 84.4%; Pred. No. 1,7e-187;
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

Qy 12 GAGGGGGCTCTCGCTGCTCGGAGAGCGCTCTTAAGCTGCTGTGCGCGGAGAGTTTG 71
Db 196 GACCGGGGCTCTTGTGCTGCTGCGGAGCGCTCTCT--AAGTACCTGTGTGCGCGGAGAGTTTG 254
Qy 72 GAGGGGCGGGCTTAGGGTCAAGTTGCGGTGGGGGGCTCGCAGCG--GACCCTCAGATCTCC 129
Db 255 GAGGGGCGGGTTTGGGGTGGGTGTGTGATTTGGGGCTCGCAGCGAGCGAGATCTCC 314
Qy 130 GCTTAGGTGCTTAGTTAAGTGTGCGGAAGCTGGGCGCAGCGCGTTC--ACTGGCCACCTGAAC 188
Db 315 GCTTAGGTACCACTTAGCGTCAAGGGAGCGTGGGTCAAGGGGGTGTGCGGAGACCCCGTGT 374
Qy 189 CTGGCGGGAGCCGAGCGCTCTGTGAGAAAGCGGGACAGCCCGTTTTCACAGCAGCTG 248
Db 375 GTGGCAGGGGCGGAAGCGCTGTGAGAAATCCCGACAGCGCTCTCTCCAGCAGGTG 434
Qy 249 CTAGGGTGGGAGCCCAAGAAAGTGAAGTGAAGTCCGGCTGCTTCCAGAGCGCTGGCC 308
Db 435 -TAGTTTCGGGAGCCCACTGGGGGCAAGTGAAGTCAAGCGGTCTTCCAGCGCTTGGGCC 493
Qy 309 ACGGCGGGCGCGGTGGGAGCAGAGTGAAGCAACCTGTTACACTAAGATGAAGGCTG 368
Db 494 ACGGCGGGCGCGCTGGGAGCAGAGTGAAGCAACCTGTTACACTAAGATGAAGGCTG 553
Qy 369 GGGTGGCTAGCCCTACTTTTGGGGGCTCTGCGGGAATGCTGCGGCTCGAAGAGCCA 428
Db 554 GGGTGGCTAGCCCTACTTTTGGGGGCTCTGCGGGAATGCTGCGGCTCGAAGAGCCA 613
Qy 429 AGATCTACACTGTGTGAGCTTTCAGAGGCTGTGTGATGAATTAAGATGGGAATTTGCCCG 488
Db 614 GGAATCTCACTGTGTGAGCAATGAGAGGCTGTGTGATGAATTAAGATGGGAATTTGCCCA 673
Qy 489 CGTGGACCCCAAGAAAGCAATTCAGATGGATCTTCCGAATCAATCAAGATGGCAGCCA 548
Db 674 GGTGGACCCCAAGAAAGCAATTCAGATGGATCTTCCGAATCAATCAAGATGGCAGCCA 733
Qy 549 GTCAGTGTGTGAGTACCTTATGCCCCCTCAGAGAGCCACCTCAGAGATGTGTGGTGA 608
Db 734 GTCAGTGTGTGAGTGTGCTTATGCCCCCTCAGAGAGCCACCTCAGAGATGTGTGGTGA 793
Qy 609 GGTGTGTGACCGAATGAGAGATGAGGGGAACAGATTGACCTTTCACCCAGCCAGAA 668
Db 794 GATATGTGACCGAATGAGAGATGAGGGGAACAGATTGATCTTCCAGCCATGGCAAGAA 853
Qy 669 CTAGTACGGGTGTGTGAGCCCGGAATGGAATCCAGTGAATGACTTACAGGGCATCCG 728
Db 854 CTAGTACGGGTGTGTGAGCCCGGAATGGAATCCAGTGAATGACTTACAGGGCATCCG 913
Qy 729 AATTGACTCAGATATCAGCGCACCCCTCAATGTTGGGTGAGAGGATGTGGTGA 788
Db 914 AATGCACTCAGATATTAAGCGCACCCCTCAATGTTGGGTGAGAGGATGTGGTGA 973
Qy 789 CGAGATGAGCTTATGCAATTTCTTCCAGAGAGCTGCAACAGTTAAAGACAACCTTG 848
Db 974 CGAGATGAGCTTATGCAATTTCTTCCAGAGAGCTGCAACAGTTAAAGACAACCTTG 1033
Qy 849 CAGTAAGCGGAGATCTATGTGACATGCCCGCACAGATCTCAGATGATGATGAAT 908
Db 1034 CAGTAAGCGGAGATCTATGTGACATGCCCGCACAGATCTCAGATGATGATGAAT 1093
Qy 909 CACTGAGCAAGCAGCTTACACCAACCTGTGATGAACACCCAGAGAGGAGATGACA 968
Db 1094 CACTGAGCAAGCAGCTTACACCAACCTGTGATGAACACCCAGAGAGGAGATGACA 1148
```

QY 969 GCATTGCTTT--TATATTACGTTTATGGAATGAAGTGAATAA 1013  
Db 1149 GCAATGCTTTATATATATATGTTTACTGAAATTAAGTGAATAA 1195

## RESULT 8

US-10-137-865-173  
: Sequence 173, Application US/10137865  
: Publication No. US20030032155A1  
: GENERAL INFORMATION:  
: APPLICANT: Baker, Kevin P.  
: APPLICANT: Beresini, Maureen  
: APPLICANT: DeForge, Laura  
: APPLICANT: Desnoyers, Luc  
: APPLICANT: Filvaroff, Ellen  
: APPLICANT: Gao, Wei-Qiang  
: APPLICANT: Gerritsen, Mary E.  
: APPLICANT: Goddard, Audrey J.  
: APPLICANT: Godowski, Paul J.  
: APPLICANT: Gurney, Austin L.  
: APPLICANT: Sherwood, Steven  
: APPLICANT: Smith, Victoria  
: APPLICANT: Stewart, Timothy A.  
: APPLICANT: Tumas, Daniel  
: APPLICANT: Watanabe, Colin K  
: APPLICANT: Wood, William  
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
: FILE REFERENCE: P330R1C154  
: CURRENT FILING DATE: 2002-05-03  
: Prior Application removed - See Palm or File Wrapper  
: NUMBER OF SEQ ID NOS: 550  
: SEQ ID NO 173  
: LENGTH: 1210  
: TYPE: DNA  
: ORGANISM: Homo Sapien  
US-10-137-865-173

Query Match 64.68; Score 691; DB 9; Length 1210;  
Best Local Similarity 84.48; Pred. No. 1.7e-187;  
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;  
QY 12 GAGGGGGTCTCTGCTGCTGAGGCGCTCTAAAGCTGCTGCTGCGGAGAGTTG 71  
Db 196 GACGGGGTCTCTGCTGCTGAGGCGCTCTCT-AGCTGCTGCTGCGGAGAGTTG 254  
QY 72 GAGGGGGGCTTACGTCAGTTCGTTGGGGGCTGCGCAGG--GACCTCAGATCTCC 129  
Db 255 GAGGGGGGCTTACGTCAGTTCGTTGGGGGCTGCGCAGGCGCAGCTGAGTCC 314  
QY 130 GCTTACGTCAGTTCGTTGGGGGCTGCGCAGGCGCTGCGCAGGCGCTGCGCAGG 188  
Db 315 GCTTACGTCAGTTCGTTGGGGGCTGCGCAGGCGCTGCGCAGGCGCTGCGCAGG 374  
QY 189 CTGCGGAGGCGGAGCGCTCTGAGAAAGCGGAGCGGAGCGGCGCTGCGCAGGCTG 248  
Db 375 GTGCGAGGCGGAGCGCTCTGAGAAAGCGGAGCGGAGCGGCGCTGCGCAGGCTG 434  
QY 249 CTAGGCTTGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 308  
Db 435 -TATGTTGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 493  
QY 309 ACGGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 368  
Db 494 ACGGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 553  
QY 369 GGGTGGCTACGCTTACTTTTGGGGGCTCTCTGGAAGCTGCTGGGCTGCAAGAGCA 428  
Db 554 GGGTGGCTACGCTTACTTTTGGGGGCTCTCTGGAAGCTGCTGGGCTGCAAGAGCA 613

QY 429 ACATACACTGTGAGCTTGCAGGCGCTCTGCTGATGAATAGTGGAAATTTCCCG 488  
Db 614 GATCTCCACTGTGAGATGAGATGAGGCGCTCTGCTGATGAATAGTGGAAATTTCCCG 673  
QY 489 GGTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 548  
Db 674 GTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 733  
QY 549 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 608  
Db 734 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 793  
QY 609 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 668  
Db 794 GATATGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 853  
QY 669 CTACGTACGCGTGTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 728  
Db 854 CTACGTACGCGTGTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 913  
QY 729 AATTGACTCAGATATACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 788  
Db 914 AATGACTCAGATATACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 973  
QY 789 CGAGATGAGCTTATGCAATTTCTTCCAGAGAGGCTGACAAGCTTAAAGCAACTT 848  
Db 974 CGAGATGAGCTTATGCAATTTCTTCCAGAGAGGCTGACAAGCTTAAAGCAACTT 1033  
QY 849 CAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 908  
Db 1034 CAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1093  
QY 909 CACTGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 968  
Db 1094 CACTGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1148  
QY 969 GCATTGCTTT--TATATTACGTTTATGGAATGAAGTGAATAA 1013  
Db 1149 GCAATGCTTTATATATATATGTTTACTGAAATTAAGTGAATAA 1195

RESULT 9  
US-10-140-474-173  
: Sequence 173, Application US/10140474  
: Publication No. US20030032156A1  
: GENERAL INFORMATION:  
: APPLICANT: Baker, Kevin P.  
: APPLICANT: Beresini, Maureen  
: APPLICANT: DeForge, Laura  
: APPLICANT: Desnoyers, Luc  
: APPLICANT: Filvaroff, Ellen  
: APPLICANT: Gao, Wei-Qiang  
: APPLICANT: Gerritsen, Mary E.  
: APPLICANT: Goddard, Audrey J.  
: APPLICANT: Godowski, Paul J.  
: APPLICANT: Gurney, Austin L.  
: APPLICANT: Sherwood, Steven  
: APPLICANT: Smith, Victoria  
: APPLICANT: Stewart, Timothy A.  
: APPLICANT: Tumas, Daniel  
: APPLICANT: Watanabe, Colin K  
: APPLICANT: Wood, William  
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
: FILE REFERENCE: P330R1C162  
: CURRENT FILING DATE: 2002-05-06  
: Prior Application removed - See Palm or File Wrapper  
: NUMBER OF SEQ ID NOS: 550  
: SEQ ID NO 173  
: LENGTH: 1210  
: TYPE: DNA

ORGANISM: Homo Sapien  
US-10-140-474-173

Query Match 64.6%; Score 691; DB 9; Length 1210;  
Best Local Similarity 84.4%; Pred. No. 1.7e-187;  
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

```
QY 12 GAGGGGGGCTCTGCTGCTGAGAGCGCTCTTAAAGCTGCTGCTGCGGAGAGTTG 71
DB 196 GACGGGGGCTCTGCTGCTGAGAGCTCTCT-AGCTGCTGCTGCTGCGGAGAGTTG 254
QY 72 GAGGGGGGCTTAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 129
DB 255 GAGGGGGGCTTAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 314
QY 130 GCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAA 188
DB 315 GCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAA 374
QY 189 CTGGGGGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 248
DB 375 GTGGGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 434
QY 249 CTAGGCTTGGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 308
DB 435 -TAGTTTGGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 493
QY 309 ACGGGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 368
DB 494 ACGGGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 553
QY 369 GGGTTGGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 428
DB 554 GGGTTGGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
QY 429 AGATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 488
DB 614 GATCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
QY 489 CGTGGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 548
DB 674 GGTGGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 733
QY 549 GTCAGTGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 608
DB 734 GTAGTGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
QY 609 GGTGTGTGACCGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGT 668
DB 794 GATGTGTGACCGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGT 853
QY 669 CTACGTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 728
DB 854 CTACGTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
QY 729 AATGTACCTGATATGACCGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGT 788
DB 914 AATGTACCTGATATGACCGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGT 973
QY 789 CGAGGATGACCTTATGATTTCTTCCAGAGGCTGACCAAGTGAAGGAGTGAAGGAGT 848
DB 974 CGAGGATGACCTTATGATTTCTTCCAGAGGCTGACCAAGTGAAGGAGTGAAGGAGT 1033
QY 849 CAGTAAAGCGAGATCTATGACCGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGT 908
DB 1034 CAGTAAAGCGAGATCTATGACCGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGT 1093
QY 909 CACTGGAGCAAGCAGCTACCAAGCTGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 968
DB 1094 CACTGGAGCAAGCAGCTACCAAGCTGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1148
QY 969 GCATTCCTCTT--TATATTCGCTTTTATGAGAAATGAACTGAAATAA 1013
DB 1149 GCATTCCTCTT--TATATTCGCTTTTATGAGAAATGAACTGAAATAA 1195
```

DB 1149 GCAATGCCCTTTATATATATATGCTTTTACGTAATTAAGTCAAAAAA 1195

## RESULT 10

US-10-142-431-173

Sequence 173, Application US/10142431

Publication No. US20030036179A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresford, Laureen

APPLICANT: Deforge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Olang

APPLICANT: Gerlitsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Gurney, Paul J.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P330101251

CURRENT FILING DATE: 2002-05-10

Prior Application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 550

SEQ ID NO 173

LENGTH: 1210

TYPE: DNA

ORGANISM: Homo Sapien

US-10-142-431-173

Query Match 64.6%; Score 691; DB 9; Length 1210;  
Best Local Similarity 84.4%; Pred. No. 1.7e-187;  
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

```
QY 12 GAGGGGGGCTCTGCTGCTGAGAGCGCTCTTAAAGCTGCTGCTGCTGCTGCTGCTGCT 71
DB 196 GACGGGGGCTCTGCTGCTGAGAGCTCTCT-AGCTGCTGCTGCTGCGGAGAGTTG 254
QY 72 GAGGGGGGCTTAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 129
DB 255 GAGGGGGGCTTAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 314
QY 130 GCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAA 188
DB 315 GCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAA 374
QY 189 CTGGGGGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 248
DB 375 GTGGGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 434
QY 249 CTAGGCTTGGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 308
DB 435 -TAGTTTGGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 493
QY 309 ACGGGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 368
DB 494 ACGGGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 553
QY 369 GGGTTGGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 428
DB 554 GGGTTGGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
QY 429 AGATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 488
DB 614 GATCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
```







## RESULT 12

US-10-140-002-173  
Sequence 173, Application US/10140002  
Publication No. US20030037623A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Geriltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin J.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C59  
CURRENT APPLICATION NUMBER: US/10/140,002  
CURRENT FILING DATE: 2002-05-06  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-140-002-173

## Query Match

64.6%; Score 691; DB 9; Length 1210;  
Best Local Similarity 84.4%; Pred. No. 1,7e-187;

Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

QY 12 GAGGGGGGCTCGCTCCCGGAGCGCTCTAAAGCTCCGCTCGCGGAGCTTTG 71  
DB 196 GACCGGGGCTCTGCTGCTGCGGAGCTCTCCCT-AGCTGCTGCTCGCGGAGCTTTG 254  
QY 72 GAGGGGGGCTTGAAGGCTAGTTGGTGGGGGCTCGCAGG- -GACCTCGATCTCC 129  
DB 255 GAGGGGGGCTTGGGGTGGCTGTGATTTGGGGCTCGCAGCGCAGCAGCTGAGTCC 314  
QY 130 GCTTAAAGTCTTAAAGTGTGCGGAGCTGGGCGCAGCGCGTC-ACTGGCCACCTGAAC 188  
DB 315 GCTTAAAGTCTTAAAGTGTGCGGAGCTGGGCGCAGCGCGTCGCGGAGCACCCTGT 374  
QY 189 CTGGGGGAGCGGAGCGCTCTGAGAGACCGGAGACCGCCCTTTTCCAGCCAGCTG 248  
DB 375 GTGGCAGGGGCGGAGCGCTCTGAGAGACCGGAGACCGCTCTCCCTGACCGAGGTG 434  
QY 249 CTGAGGTGGGAGCCACAGAAAAGTAGAGTCCGGGTGTTCCAGAGCGCTGGGCG 308  
DB 435 -TAGTTTCGGAGACCACTGGGGCGCAAGTAGAGTCCAGCGGCTTCCAGCGCTTGGGCG 493  
QY 309 ACGGCGGGCGGCTGGAGCAGAGTGTGAGCGACCTGTATACCTAAAGATGAAGCTG 368  
DB 494 ACGGCGGGCGGCTGGAGCAGAGTGTGAGCGACCGCCTTAAAGTGAAGAGCTG 553  
QY 369 GGGTGGCTAGGCTTAAAGTGTGGGGTCCGCTGAGGAGTCCCTGGGCTCGGAAGAGCCA 428  
DB 554 GGGTGGCTAGGCTTAAAGTGTGGGGTCCGCTGAGGAGTCCCTGGGCTCGGAAGAGCCA 613  
QY 429 AGATCTACACTGTGAGCTGTGAGGCTGTGTGATGATTAAGATGGAAATTCGCG 488  
DB 614 GATCTACACTGTGAGCTGTGAGGCTGTGTGATGATTAAGATGGAAATTCGCG 673  
QY 489 CGTGAAGCCCAAGAGACCATTCAGATGGATCTCCGAATCAATCCAGATGGAGCCA 548  
|||||

## RESULT 13

US-10-142-419-173

Sequence 173, Application US/10142419  
Publication No. US20030044945A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Geriltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin J.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C244  
CURRENT APPLICATION NUMBER: US/10/142,419  
CURRENT FILING DATE: 2002-05-10  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-142-419-173

## Query Match

64.6%; Score 691; DB 9; Length 1210;  
Best Local Similarity 84.4%; Pred. No. 1,7e-187;

Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

DB 674 GGTGAGCCCAAGAGACCATTCAGATGGAGTCTTCCGATCAATCCAGATGGAGCCA 733  
QY 549 GTGAGTGTGAGGATACCTTATGCGCGCTCAGAGCGCCCACTTACAGATGCTTGAGCA 608  
DB 734 GTGAGTGTGAGGATACCTTATGCGCGCTCAGAGCGCCCACTTACAGATGCTTGAGCA 793  
QY 609 GGTGTGTGAGCCGATGAGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 668  
DB 794 GATATGTGAGCCGATGAGAGGATGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 853  
QY 669 CTACGTACGCTGTGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 728  
DB 854 CTACGTACGCTGTGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 913  
QY 729 AATTGACTGATATGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 788  
DB 914 AATCGACTGATATGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 973  
QY 789 CGAGATGAGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 848  
DB 974 CGAGATGAGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1033  
QY 849 CAGTAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 908  
DB 1034 CAGTAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1093  
QY 909 CACTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 968  
DB 1094 CACTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1148  
QY 969 GCATGCTCTT-TATATTAGTTTATGAGAAATGAGTGAAGAAA 1013  
DB 1149 GCATGCTCTTATATATATATGTTTACTGAGAAATGAGTGAAGAAA 1195  
|||||

Query Match 64.6%; Score 691; DB 9; Length 1210;  
Best Local Similarity 84.4%; Pred. No. 1,7e-187;  
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;





Country	1950	1960	1970	1980	1990	2000	2010	2020	2030	2040	2050
Japan	7	8	10	12	14	16	18	20	22	24	26
Germany	10	11	12	13	14	15	16	17	18	19	20
France	11	12	13	14	15	16	17	18	19	20	21
Italy	12	13	14	15	16	17	18	19	20	21	22
Spain	13	14	15	16	17	18	19	20	21	22	23
Sweden	14	15	16	17	18	19	20	21	22	23	24
United Kingdom	15	16	17	18	19	20	21	22	23	24	25
United States	16	17	18	19	20	21	22	23	24	25	26
Canada	17	18	19	20	21	22	23	24	25	26	27
Belgium	18	19	20	21	22	23	24	25	26	27	28
Netherlands	19	20	21	22	23	24	25	26	27	28	29
Australia	20	21	22	23	24	25	26	27	28	29	30
South Korea	21	22	23	24	25	26	27	28	29	30	31
India	22	23	24	25	26	27	28	29	30	31	32
China	23	24	25	26	27	28	29	30	31	32	33
South Africa	24	25	26	27	28	29	30	31	32	33	34
Brazil	25	26	27	28	29	30	31	32	33	34	35
Argentina	26	27	28	29	30	31	32	33	34	35	36
Colombia	27	28	29	30	31	32	33	34	35	36	37
Peru	28	29	30	31	32	33	34	35	36	37	38
Venezuela	29	30	31	32	33	34	35	36	37	38	39
Chile	30	31	32	33	34	35	36	37	38	39	40
Ecuador	31	32	33	34	35	36	37	38	39	40	41
Bolivia	32	33	34	35	36	37	38	39	40	41	42
Paraguay	33	34	35	36	37	38	39	40	41	42	43
Uruguay	34	35	36	37	38	39	40	41	42	43	44
Costa Rica	35	36	37	38	39	40	41	42	43	44	45
Panama	36	37	38	39	40	41	42	43	44	45	46
Dominican Republic	37	38	39	40	41	42	43	44	45	46	47
Honduras	38	39	40	41	42	43	44	45	46	47	48
Guatemala	39	40	41	42	43	44	45	46	47	48	49
El Salvador	40	41	42	43	44	45	46	47	48	49	50
Nicaragua	41	42	43	44	45	46	47	48	49	50	51
Haiti	42	43	44	45	46	47	48	49	50	51	52
Dominican Republic	43	44	45	46	47	48	49	50	51	52	53
Jamaica	44	45	46	47	48	49	50	51	52	53	54
Trinidad and Tobago	45	46	47	48	49	50	51	52	53	54	55
Barbados	46	47	48	49							

10

10

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 18:36:03 ; Search time 1522.26 Seconds

(without alignments)  
11373.235 Million cell updates/sec

Title: US-10-082-502-18

Perfect score: 1069  
Sequence: 1 gaattcgacagaggggggtc.....aaaaaaattccgcgcgcgc 1069

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:\*  
1: em\_estba:\*  
2: em\_estchum:\*  
3: em\_estcin:\*  
4: em\_estcinu:\*  
5: em\_estcov:\*  
6: em\_estcpl:\*  
7: em\_estcro:\*  
8: em\_estc1:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hrc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pin:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	889.4	83.2	1255	11 AK019927	Mus muscu
2	852	79.7	853	11 AK013014	Mus muscu
3	835.4	78.1	992	13 B1456021	603170466
4	817.2	76.4	856	13 B1454180	603170602
5	813.8	76.1	891	13 B1408381	602864013
6	771	72.1	945	13 B1412341	602991837

7	770	72.0	772	11 AK007914	Mus muscu
8	745.8	69.8	754	11 AK013358	Mus muscu
9	744.6	69.7	802	12 BF608664	MY1_00160
10	734	68.7	853	12 BF535071	602054331
11	721	67.4	808	13 B1408525	B1408525
12	712.8	66.7	863	13 B1454899	B1454899
13	708.8	66.3	843	13 B1412533	B1412533
14	686	64.2	709	13 B1411532	B1411532
15	654.4	61.2	894	14 BQ963382	BQ963382
16	641	60.0	663	13 B108692	B108692
17	635.4	59.4	888	13 B1412220	602990252
18	634	59.3	1199	12 BF539304	BF539304
19	627.8	58.7	724	14 BQ210416	BQ210416
20	619.6	58.0	867	13 BF973691	BF973691
21	612.6	57.3	731	12 BF532558	BF532558
22	612.2	57.3	1117	13 BM552907	BM552907
23	603.4	56.4	875	13 B1763950	B1763950
24	600.6	56.2	609	10 AM825757	AM825757
25	599.8	56.1	604	10 BE308275	BE308275
26	599.8	56.1	880	14 BQ421827	BQ421827
27	591.4	55.3	678	10 BE309953	BE309953
28	591.4	55.3	791	12 BG704443	BG704443
29	591.4	55.3	914	12 BE799495	BE799495
30	590.8	55.3	594	10 AM908970	AM908970
31	590.6	55.2	868	14 BQ421069	BQ421069
32	585.8	54.8	887	13 B1762759	B1762759
33	585.6	54.8	600	10 AM413913	AM413913
34	582.8	54.5	596	10 BE134206	BE134206
35	582.2	54.5	783	13 B1550436	B1550436
36	581.4	54.4	595	10 AM414039	AM414039
37	576.4	53.9	742	12 BG824549	BG824549
38	573.2	53.6	656	12 BF287107	BF287107
39	571	53.4	783	13 B1763784	B1763784
40	568.4	53.2	570	10 AM826122	AM826122
41	567.8	53.1	607	12 BE956857	BE956857
42	567.4	53.0	697	12 BE344334	BE344334
43	566.2	53.0	657	14 BQ211867	BQ211867
44	563.8	52.7	665	10 AM919569	AM919569
45	558	52.2	776	13 B1522523	B1522523

#### ALIGNMENTS

RESULT 1	AK019927	1255 bp	mus musculus adult male pituitary gland cDNA, RIKEN full-length enriched library, clone:530432A10;transmembrane protein 4, full insert sequence.
LOCUS	AK019927		
DEFINITION	AK019927		
ACCESSION	AK019927		
VERSION	AK019927.1		
KEYWORDS	HTC; CAP		
SOURCE	Mus musculus (strain: C57BL/6J) adult male pituitary gland cDNA to mRNA, clone: lib:RIKEN full-length enriched mouse cDNA library		
ORGANISM	Mus musculus		
REFERENCE	1		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		



Db 483 AGGAGTACGGGGAACAGATTGACCTTCTACCCACCCGCAAGAACTACGTACGCTGTGA 542  
 Oy 686 GCCGGAATGAGAAATCCAGTGAACCTAGACCTTACAGGGCATCCGAATTGACTCAGATATCA 745  
 Db 543 GCCGGAATGAGAAATCCAGTGAACCTAGACCTTACAGGGCATCCGAATTGACTCAGATATCA 602  
 Oy 746 GCGGCAACCCCAAGTTGGCTGTGAGAGCAATTGTGGAGAAATACGAGATGAGCTTATTCG 805  
 Db 603 GCGGCAACCCCAAGTTGGCTGTGAGAGCAATTGTGGAGAAATACGAGATGAGCTTATTCG 662  
 Oy 806 AATTCTTCTCAGAGAGGCTGACAGACGTTAAAGACAACTTTGAGTAAGCGGAGACAT 865  
 Db 663 AATTCTTCTCAGAGAGGCTGACAGACGTTAAAGACAACTTTGAGTAAGCGGAGACAT 722  
 Oy 866 TATGTACCATGCTCCCTGACAGATCTCAGATGAGTATGATCACTGAGCAAGCAGCC 925  
 Db 723 TATGTACCATGCTCCCTGACAGATCTCAGATGAGTATGATCACTGAGCAAGCAGCC 782  
 Oy 926 TACACCAACGATGAGAACACCCCAAGGAGGAGGAGATGGCAGCATGCTTTTATATT 985  
 Db 783 TACACCAACGATGAGAACACCCCAAGGAGGAGGAGATGGCAGCATGCTTTTATATT 842  
 Oy 986 ACCTTTTATGAAATGAATGAAAAAACTCTTGAACCGAAGTAAAAA 1036  
 Db 843 ACCTTTTATGAAATGAATGAAAAAACTCTTGAACCGAAGTAAAAA 893

RESULT 2  
 AK013014 853 bp mRNA linear HTC 19-JAN-2002  
 LOCUS Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length  
 DEFINITION enriched library, clone:2810406L24:transmembrane protein 4, full  
 insert sequence.

ACCESSION AK013014.1 GI:12850114  
 VERSION AK013014.1  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (strain:C57BL/6J) 10, 11 days embryo cDNA to mRNA,  
 clone:lib:RIKEN full-length enriched mouse cDNA library  
 clone:2810406L24.

## ORGANISM

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 AUTHORS Carninci, P., and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636

## REFERENCE

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159

## REFERENCE

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Kono, H., Akiyama, J., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M.,  
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K.,  
 Yoneoka, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multichannel sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861

## REFERENCE

AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
 Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Aichi, T., Fukuda, S.,  
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I.,

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

## TITLE

JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

## COMMENT

COMMENT  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

## FEATURES

source  
 gene  
 CDS  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM:DB:2810406L24"  
 /db\_xref="MGI:1892961"  
 /db\_xref="taxon:10090"  
 /clone="2810406L24"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="10, 11 days embryo"  
 1. 853  
 /gene="Tmem4"  
 214. 762  
 /gene="Tmem4"

Query Match	Best Local Similarity	79.7%	Score 852;	DB 11;	Length 853;
Matches 852;	Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
146	AGATCGGGGAAGCTGGGGGCGAGGGCGGTCACTGCGCCACCCCTGAACCTGGCGGGAGCGCGGAGC	205			
2	AAGTCGCGGAAGCTGGGCGAGCGGTCTACTGCGCCACCTGAACCTGGCGGGAGCGCGGAGC	61			
206	GCTCTGGAAGACCGCGGACAGCCCGCTTTTCCAGCAGCTGCTAGAGGTTGGACCCAC	265			
62	GCTCTGGAAGACCGCGGAGCACCCCGCTTTTCCAGCAGCTGCTAGAGGTTGGACCCAC	121			
266	AGAAAACAAGTGAAGTTCGGGCTGCTTTCCAGAGCCCTGGGCGACGCGCGCGCGCTGGG	325			
122	AGAAAACAAGTGAAGTTCGGGCTGCTTTCCAGAGCCCTGGGCGACGCGCGCGCGCTGGG	181			
326	AGCAGAGGTGAGGAGCCCTGTACACTAAAGATGAAGAGCTGGGGTGGCTAGCCCTAC	385			
182	AGCAGAGGTGAGGAGCCCTGTACACTAAAGATGAAGAGCTGGGGTGGCTAGCCCTAC	241			
386	TTTTGGGGGCTCGCTGGGAAGCTGCGTGGGCTCGAAGAGCCAAAGATTAACACTGTGGAG	445			
242	TTTTGGGGGCTCGCTGGGAAGCTGCGTGGGCTCGAAGAGCCAAAGATTAACACTGTGGAG	301			
446	CTTGCAGGGCTCTGCTGATGAATTAGAGTGGAAATTTGCCCGCGTGAGCCCAAGAAGA	505			
302	CTTGCAGGGCTCTGCTGATGAATTAGAGTGGAAATTTGCCCGCGTGAGCCCAAGAAGA	361			
506	CCATTCAATGAGGATCCCTCCGAATCAATCCAGATGGAGCAGCACTTTTGGAGGTAC	565			
362	CCATTCAATGAGGATCCCTCCGAATCAATCCAGATGGAGCAGCACTTTTGGAGGTAC	421			
566	CTTATGCCGCTCGAGAGGCCACCTCAAGAGATTGCTTGAAGAGGTGTGACCGAATGA	625			
422	CTTATGCCGCTCGAGAGGCCACCTCAAGAGATTGCTTGAAGAGGTGTGACCGAATGA	481			
626	AGCAGTACGGGGAACAGATTGATGACCCCTTCTACCCACCGCAAGACTACGTACGCGTGTGA	685			
482	AGCAGTACGGGGAACAGATTGATGACCCCTTCTACCCACCGCAAGACTACGTACGCGTGTGA	541			
686	GCCGGAATGAGGAATCCAGTGAATCACTACAGAGGATCCGAATTGACATCAATATCA	745			
542	GCCGGAATGAGGAATCCAGTGAATCACTACAGAGGATCCGAATTGACATCAATATCA	601			
746	GCGGACCCCTCAAGTTTACGTGTGAGACATTTGGGAAGATTAAGAGATAGCTTATCG	805			
602	GCGGACCCCTCAAGTTTACGTGTGAGACATTTGGGAAGATTAAGAGATAGCTTATCG	661			
806	AATTCCTCTCCAGAGAGGCTGACAACTTAAGAACTTTGCAAGTAAAGCGACAGATTC	865			
662	AATTCCTCTCCAGAGAGGCTGACAACTTAAGAACTTTGCAAGTAAAGCGACAGATTC	721			
866	TATGTGACATGCTCCGTGACAGATCTCAGATGAGCTATGATTAATCACTGAGACAAAGACC	925			
722	TATGTGACATGCTCCGTGACAGATCTCAGATGAGCTATGATTAATCACTGAGACAAAGACC	781			
926	TATGCAAAAGCTGATGGAAACACCCCGAGAGGGGAAGATGGCAGCATTTGCTTTATATT	985			

Db	782	TACACCAAAAGCTGATGGAACACCCCGAGAGGGGGAAGATGGCAGCATGCTTTATATT	841
QY	986	ACGTTTATGCG	997
Db	842	ACGTTTATGCG	853
RESULT 3			
BI456021			
LOCUS	BI456021	992 bp	mRNA linear EST 21-AUG-2001
DEFINITION	603170466F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5249992 5'		
ACCESSION	BI456021		
VERSION	BI456021.1	GI:15246677	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	1 (bases 1 to 992)		
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgaps-remail.nih.gov">cgaps-remail.nih.gov</a> Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) DNA Sequencing by: Incyte Genomics, Inc. DNA distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LHM11631 Row: C Column: 17 High quality sequence stop: 854. Location/Qualifiers 1..992 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:5249992" /clone_id="NCI_CGAP_Mam5" /tissue_type="tumor, gross tissue" /dev_stage="7 months" /lab_host="DH10B" /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"		
BASE COUNT	256 a	246 c	295 g 194 t 1 others
ORIGIN			
Query Match	78.1%	Score 835.4;	DB 13; Length 992;
Best Local Similarity	97.4%	Pred. No. 2e-123;	
Matches 923; Conservative	0;	Mismatches 17;	Indels 8; Gaps 7;
QY	103	GGGCTCGACGAGGACCTCAGATCTCGCTTAGTGTGCTTAAATGTCGGGAAGCTGGG	162
Db	6	GCGNTCGAGCGAGACCTCAGATCTCGCTTAGTGTGCTTAAATGTCGGGAAGCTGGG	65
QY	163	CGAAGCGGTACTGTGGCCACCTGAACTCGGCGGAGCCGGAGCGCTCTGAGAAACCGGG	222
Db	66	CGAAGCGGTACTGTGGCCACCTGAACTCGGCGGAGCCGGAGCGCTCTGAGAAACCGGG	125
QY	223	ACAGCCCGGTTTTCCAGCGACGCTGCTGAAGGCTTGGGACCCACAGAAACAAAGTGAGAG	282
Db	126	ACAGCCCGGTTTTCCAGCGACGCTGCTGAAGGCTTGGGACCCACAGAAACAAAGTGAGAG	185
QY	283	TCCGCGTGTCTTCAGAGACCTGGGCGACAGCGGCGGCGCGTGGAGCAGAGGTGAGCGAC	342
Db	186	TCCGCGTGTCTTCAGAGACCTGGGCGACAGCGGCGGCGCGTGGAGCAGAGGTGAGCGAC	245
QY	343	CCGTGTACACTAAAGATGAAGGCTGGGGTTGGCTAGCCCTACTTTTGGGGGCTCTGCTG	402



DB 246 CCTGTACCTTAAGTGAAGGCTGGGTTGGCTAGCCCTTCTTTGGGGTCTCTG 305  
OY 403 GGAAGTGGCTGGGCTGAAGAGCCAGATCTACACTGTGGAGCTTGAGGGCTGTG 462  
DB 306 GGAAGTGGCTGGGCTGAAGAGCCAGATCTACACTGTGGAGCTTGAGGGCTGTG 365  
OY 463 GATGAATTAAGATGGGAAATTTCCCGCGTGGAGCCCAAGAACCAATTCAGATGGATCC 522  
DB 366 GATGAATTAAGATGGGAAATTTCCCGCGTGGAGCCCAAGAACCAATTCAGATGGATCC 425  
OY 523 TTCCCAATCAATCCAGATGCGAGCCAGTCACTGATGGAGTCACTTATGCCGCTCAGAG 582  
DB 426 TTCCCAATCAATCCAGATGCGAGCCAGTCACTGATGGAGTCACTTATGCCGCTCAGAG 485  
OY 563 GCCACCTTCACAGAGTGGTGGAGAGGTGTGTGACCCGAATGAGAGAGTACGGGGAGACAG 642  
DB 486 GCCACCTTCACAGAGTGGTGGAGAGGTGTGTGACCCGAATGAGAGAGTACGGGGAGACAG 545  
OY 643 ATTGACCTTCTACCCACCCGCAAGACTACGCTGCTGTAGCCGGAATGAGAGATCC 702  
DB 546 ATTGACCTTCTACCCACCCGCAAGACTACGCTGCTGTAGCCGGAATGAGAGATCC 605  
OY 703 AGTGAAGTGAAGTGAAGGAGCATCCGAATTTGACTCAGATATCAGGGGCACTCAAGTTT 762  
DB 606 AGTGAAGTGAAGTGAAGGAGCATCCGAATTTGACTCAGATATCAGGGGCACTCAAGTTT 665  
OY 763 GCGTGTGAGAGCATTTGTGGAAGATACGAGATGAGCTTATGAAATTTCTTCACAGAGAG 822  
DB 666 GCGTGTGAGAGCATTTGTGGAAGATACGAGATGAGCTTATGAAATTTCTTCACAGAGAG 725  
OY 823 GCTGACAA--CCTTAAAGACAACTTTGCAATGAGGAGCAGATCT--ATGACACATGCC 879  
DB 726 GCTGACAAACGTTTAAAGACAACTTTGCAATGAGGAGCAGATCTTAAATGACATGTCG 785  
OY 880 CTGCACAGATCTCAGATGAGC--TATGAATCACT--GAGACAGACAGCCATCAACCAAGCT 937  
DB 786 CTGCACAGATCTCAGATGAGTAACTTTGAAATCACTGGAGACAGAGTCAAC--AAGCT 844  
OY 938 GATGAGACACCCCAAGAG--GGGGAAGATGAGCAGATGCTTATATATACCTTTTATG 996  
DB 845 GATGAGACACCCCAAGAGAGGAGGAGATGAGCAGATGCTTATATATACCTTTTATG 904  
OY 997 GAAATGAATGAAAAAACTCTTGAACCGGAATGAAAAAAATTTT 1044  
DB 905 GGATTAAGCTG--AAAAAATCTTTGAACCGGAATGAAAAAGACCAAAA 951

RESULT 4  
BI454180 856 bp mRNA linear EST 21-AUG-2001  
LOCUS DEFINITION 603170602P1 NCI\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:5250001 5',  
ACCESSION BI454180  
VERSION BI454180.1 GI:15244836  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 856)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-r@mail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNM)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNM at:  
http://image.llnl.gov

Plate: LLNM11631 row: d column: 02  
High quality sequence stop: 826.  
FEATURES  
Location/Qualifiers  
1..856  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5250001"  
/clone\_id="NCI\_CGAP\_Mam5"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-Sport6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NIH"  
BASE COUNT 207 a 217 c 266 g 166 t  
ORIGIN  
Query Match 76.4%; Score 817.2; DB 13; Length 856;  
Best Local Similarity 99.3%; Pred. No. 1.7e-120;  
Matches 852; Conservative 0; Mismatches 3; Indels 3; Gaps 3;  
OY 93 TTTCCGTTGGGGGCTGCGACGGGACCTCAGATCTCCGTTAGTGGCTTAAAGTGGC 152  
DB 1 TTTCCGTTGGGGGCTGCGACGGGACCTCAGATCTCCGTTAGTGGCTTAAAGTGGC 60  
OY 153 GGAAGTGGGGGCGGCGGCTGCGACGGGACCTCAGATCTCCGTTAGTGGCTTAAAGTGGC 212  
DB 61 GGAAGTGGGGGCGGCGGCTGCGACGGGACCTCAGATCTCCGTTAGTGGCTTAAAGTGGC 120  
OY 213 AGAAGCGGGAGACGCCCGCTTTTCCAGCAGCTGCTAGGCTTGGGACCCAGAAAAC 272  
DB 121 AGAAGCGGGAGACGCCCGCTTTTCCAGCAGCTGCTAGGCTTGGGACCCAGAAAAC 180  
OY 273 AAAGTGAAGTCCGGCTCTTTCAGAGCTGCGCCACGGCGGCGCTGGAGACAGAG 332  
DB 181 AAAGTGAAGTCCGGCTCTTTCAGAGCTGCGCCACGGCGGCGGCTGGAGACAGAG 239  
OY 333 GTGGAGCGACCGCTGTACCTAAAGATGAAGGCTGGGCTTGGCTAGGCTTCTTTGGG 392  
DB 240 GTGGAGCGACCGCTGTACCTAAAGATGAAGGCTGGGCTTGGCTAGGCTTCTTTGGG 299  
OY 393 GGTCTGCTGGGAGACCTGCTGGGCTGGAAGAGCCAAAGATCTACACTGTGAGCTTGAC 452  
DB 300 GGTCTGCTGGGAGACCTGCTGGGCTGGAAGAGCCAAAGATCTACACTGTGAGCTTGAC 359  
OY 453 GGTCTGCTGGGAGATGAATTAAGTGGGAAATTTCCCGCTGGAACCCCAAGACCATTC 512  
DB 360 GGTCTGCTGGGAGATGAATTAAGTGGGAAATTTCCCGCTGGAACCCCAAGACCATTC 419  
OY 513 GATGGAGTCCCTCCGAATATCCAGATGGAGCCAGCTAGTTGGAGAGTACCTTATGC 572  
DB 420 GATGGAGTCCCTCCGAATATCCAGATGGAGCCAGCTAGTTGGAGAGTACCTTATGC 479  
OY 573 CCGCTCAGAGGCCCACTTCACAGAGTGGCTTGAAGAGGTGTGTGACCAATGAAGAGTA 632  
DB 480 CCGCTCAGAGGCCCACTTCACAGAGTGGCTTGAAGAGGTGTGTGACCAATGAAGAGTA 539  
OY 633 CCGGGAACAGATTGACCTTCTACCCACCGCAAGAACTACGTAACGGCTGTAGCCGGA 692  
DB 540 CCGGGAACAGATTGACCTTCTACCCACCGCAAGAACTACGTAACGGCTGTAGCCGGA 599  
OY 693 TGGAGAACTCAATGAACTGACTTACAGGGCATCCGAATTTGACTCAGATATCAGGGCAC 752  
DB 600 TGGAGAACTCAATGAACTGACTTACAGGGCATCCGAATTTGACTCAGATATCAGGGCAC 659  
OY 753 CCTCAAGTTGGGTGTGAGAGCATTTGTGGAAGAAATACGAGATGAGCTTATGAAATCTT 812  
DB 660 CCTCAAGTTGGGTGTGAGAGCATTTGTGGAAGAAATACGAGATGAGCTTATGAAATCTT 719  
OY 813 CTCAGAGAGGCTGACAACTTAAAGACAACTTTGCAAGTAAAGCGAGACGATCTATGTGA 872

[illegible]

OY	294	TCACAGACCTGGGGACACGGCGGGCCGCGGGAGAGAGAGTGGAGGACCCGTTTACACT	353
Db	134	TCCAGAGCCTGGGGCCACGGCGGGCCGCGGGAGAGAGTGGAGGACCCGTTTACACT	193
OY	354	AAAGATGAAGAGCTGGGGTGTGGCTAGCCCTACTTTTGGGGGTCCTGCTGGAACTGCTG	413
Db	194	AAAGATGAAGAGCTGGGGTGTGGCTAGCCCTACTTTTGGGGGTCCTGCTGGAACTGCTG	253
OY	414	GGCTGGAGGAGACCAAGATCTACACTGTGTGAGACTTTCAGAGGCTCTGTGTGATGAATTAGA	473
Db	254	GGCTGGAGGAGACCAAGATCTACACTGTGTGAGACTTTCAGAGGCTCTGTGTGATGAATTAGA	313
OY	474	GGGGGAAATGGCCGCGTGGAGCCCAAGAGAACCAATTGAGTGGGATGCTTCCGTAATCA	533
Db	314	GTGGGAAATGGCCGCGGTGGAGCCCAAGAGAACCAATTGAGTGGGATGCTTCCGTAATCA	373
OY	534	TCCAGATGGCAGCCACGTCAGTGTGTGAGGTACTTAATGCCCGCTCAGAGAGCCACCTCAC	593
Db	374	TCCAGATGGCAGCCACGTCAGTGTGTGAGGTACTTAATGCCCGCTCAGAGAGCCACCTCAC	433
OY	594	AGAGTTGCTTGGAGAGGTGTGTGACCGAATGAAGAGTACGGGGAAACAGATTGACCTTC	653
Db	434	AGAGTTGCTTGGAGAGGTGTGTGACCGAATGAAGAGTACGGGGAAACAGATTGACCTTC	493
OY	654	TACCCACGCCAAGAACTACGTACGGCGTGTGACCGCGGAATGGAGATCAGTGAATAGA	713
Db	494	TACCCACGCCAAGAACTACGTACGGCGTGTGACCGCGGAATGGAGATCAGTGAATAGA	553
OY	714	CTTTACAGGGGCATCCGAATTGACTCAGATATACAGCGGACACCTTCAAGTTTGC - TGTAGA	772
Db	554	CTTTACAGGGGCATCCGAATTGACTCAGATATACAGCGGACACCTTCAAGTTTGC - TGTAGA	613
OY	773	GCATTGTGGAAATACGAGGATGAGCTATGCAATTCCTTCGAGAGGCGTGCACACG	832
Db	614	GCATTGTGGAAATACGAGGATGAGCTATGCAATTCCTTCGAGAGGCGTGCACACG	673
OY	833	TTTAAAGACAACTTTGACAGTAAGCGGACAGATCTATGTGACCATGCCCCGACAGATCTC	892
Db	674	TTTAAAGACAACTTTGACAGTAAGCGGACAGATCTATGTGACCATGCCCCGACAGATCTC	733
OY	893	ACGATGAGCTATGAATCAGTGGAGACAGACGCTTACACCAACGATGAGGAACACCCCA	952
Db	734	ACGATGAGCTATGAATCAGTGGAGACAGACGCTTACAC - AACGTGATGGAACACCCCA	792
OY	953	GGAGGGG - AAGATGGACGACTTGGCTTTATATACGTTTTATGGAATGAACGTGAA	1010
Db	793	GGAGGGGACAGCATGGACGACTTGGCTTTATATATACGTTTTATGGAATGAACGTGAA	852
OY	1011	AAACCTTTGAACCGAAGATTAATAAAAAAAAAAAAAAAAAA	1049
Db	853	AAACCTTTGAACCGAAGATTAATAAAAAAAAAAAAAAAAAA	891
RESULT 6			
LOCUS	Bi412341	945 bp	mRNA
DEFINITION	60299183371 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5147624 5'		linear EST 14-AUG-2001
ACCESSION	Bi412341		
VERSION	Bi412341.1	GI:15173264	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	Nih-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cga@bms.fremail.nih.gov		
	Tissue Procurement: Gilbert Smith, Ph.D.		





[illegible]

	TITLE	Journal					
	JOURNAL	Kato H., Kawai J., Kojima T., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numezaki R., Ohno M., Okazaki Y., Okido T., Owa C., Quackenbush J., Salto R., Salto R., Sakai C., Sakai K., Sano H., Sasaki D., Schiml L., Shibata K., Shibata Y., Shinagawa A., Shiraki T., Sojabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Teijima Y., Toya T., Yamamura T., Yamanaka I., Yasunishi A., Yoshida K., Yoshino M., Muramatsu M. and Hayashizaki Y.					
	DIRECT SUBMISSION	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-2 Tsukuba-shi, Tsukuba-city, Yokohama Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)					
	COMMENT	Please visit our web site ( <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> ) for further details. CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGATCCACGATCCAGAGCGCTTTTGTATTATTTTAAVN 3'], cDNA was prepared by using triazole thermo-activated reverse transcriptase and subsequently amplified for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGGATTAATAATTAATCAATCCCCCCCCC 3']. cDNA was cleaved with XhoI and SacI. Cloning sites, 5' end: XhoI; 3' end: SacI. Host: SOLR.					
	FEATURES						
	SOURCE	Location/Qualifiers 1..754 /organism="Mus musculus" /strain="C57BL/6Jr" /db_xref="PANTOM.DB:2900022B12" /db_xref="MGD:MGI:1896245" /db_xref="taxon:10090" /clone="2900022B12" /sex="male" /tissue_type="hippocampus" /clone_id="RIKEN full-length enriched mouse cdna library." /des_tag="adult" 1..754 /gene="Tmem4" 85..633 /gene="Tmem4" /note="data source:MGD, source key:MGI:1928477, evidence:ISS putative transmembrane protein 4" /codon_start=1 /protein_id="BAB28909.1" /db_xref="GI:12850980" /db_xref="MGD:MGI:1928477" /translaton="MKNGMALLLGVLTGMARRBSODLHGCRCALVDELEFIAR VDPKRTIOMGSFRINPDGOSVVEVPARSABHTELLEVCDDMKETGEIDISTRRHR KNIVRVSRNSESSELDLOGRIDSISGLTFACESIVEYEDELTEFESREADNVK DKLSKRSTDCLDHLLHSHDEL"					
	CDS						
	gene						
	BASE COUNT	200 a 175 c 220 g 159 t					
	ORIGIN						
	Query Match	69.8% Score 745.8; DB 11; Length 754;					
	Best Local Similarity	99.7% Pred. No. 3.7e-109;					
	Matches 747; Conservative	0; Mismatches 2; Indels 0; Gaps 0;					
OY	278	GAGACTCCGGCTGCTTCCAGAGCCTGGGCCACGGCGCGGCGGAGCAGAGTGGA	337				
Db	5	GAGAGTCCGGGCTGCTTCCAGAGCCTGGGCCACGGCGCGGCGGAGCAGAGTGGA	64				
OY	338	GCGACCCTGTTACACTAAGAATGAAGGCTGGGGTGGCTAGCCCTACTTTTGGGGGTCC	397				
Db	65	GCGACCCTGTTACACTAAGAATGAAGGCTGGGGTGGCTAGCCCTACTTTTGGGGGTCC	124				

QY 398 TCGTGGAACTGCTTGGGCTCGAAGAGCCAAAGATCTACACTGTGGAGCTTGCAGGGCTC 457  
 |||||||  
 Db 125 TCGTGGAACTGCTTGGGCTCGAAGAGCCAAAGATCTACACTGTGGAGCTTGCAGGGCTC 184  
 |||||||  
 QY 458 TGGTGATGATTTAGTGGGAAATGGCCCGGTGGAGCCCAAGAGACCAATTCAGATGG 517  
 |||||||  
 Db 185 TGGTGATGATTTAGTGGGAAATGGCCCGGTGGAGCCCAAGAGACCAATTCAGATGG 244  
 |||||||  
 QY 518 GATCCTTCGGAATCAATCAGATGGCAGCCAGTCAGTTGTGGAGTACCTTATGCCCGCT 577  
 |||||||  
 Db 245 GATCCTTCGGAATCAATCAGATGGCAGCCAGTCAGTTGTGGAGTACCTTATGCCCGCT 304  
 |||||||  
 QY 578 CAGAGGCCCTCAGAGAGTTGCTTGGAGAGGTGTGTGACCGAATGAAGAGTACGGGG 637  
 |||||||  
 Db 305 CAGAGGCCCTCAGAGAGTTGCTTGGAGAGGTGTGTGACCGAATGAAGAGTACGGGG 364  
 |||||||  
 QY 638 AACAGATTGACCTTCTCCACCCAGCAAGTACGTACGGTGTGAGCCGGAATGGAG 697  
 |||||||  
 Db 365 AACAGATTGACCTTCTCCACCCAGCAAGTACGTACGGTGTGAGCCGGAATGGAG 424  
 |||||||  
 QY 698 AATCCAGTAACTAGACTTACAGGGCATCCGAATTGACTAGATATTCAGCGGCACTCA 757  
 |||||||  
 Db 425 AATCCAGTAACTAGACTTACAGGGCATCCGAATTGACTAGATATTCAGCGGCACTCA 484  
 |||||||  
 QY 758 AGTTGGCTGTGAGAGCATTTGTGGAAGAATACAGAGATGAGCTTATCCAAATTTCTTCCA 817  
 |||||||  
 Db 485 AGTTGGCTGTGAGAGCATTTGTGGAAGAATACAGAGATGAGCTTATCCAAATTTCTTCCA 544  
 |||||||  
 QY 818 GAGAGGCTGACAGCTTAAAGACAAATTTGACAGTAAAGCAGACAGATATTCAGCAGT 877  
 |||||||  
 Db 545 GAGAGGCTGACAGCTTAAAGACAAATTTGACAGTAAAGCAGACAGATATTCAGCAGT 604  
 |||||||  
 QY 878 CCTGACACAGATCTCAGATGAGTATGATCACTGAGCAGCAGCTTACACCAAGCT 937  
 |||||||  
 Db 605 CCTGACACAGATCTCAGATGAGTATGATCACTGAGCAGCAGCTTACACCAAGCT 664  
 |||||||  
 QY 938 GATGAACACCCCGAGAGGGGGAAGATGGCAGCATGCTTTTATATACGTTTTTATAG 997  
 |||||||  
 Db 665 GATGAACACCCCGAGAGGGGGAAGATGGCAGCATGCTTTTATATACGTTTTTATAG 724  
 |||||||  
 QY 998 AATGAACTGAAAAAACTCTTGAACCG 1026  
 |||||||  
 Db 725 AATGAACTGAAAAAACTCTTGAACCG 753  
 |||||||

RESULT 9 802 bp mRNA linear EST 01-APR-2001  
 BF608664  
 LOCUS  
 DEFINITION  
 CDNA clone ICRP522B2440 5', mRNA sequence.  
 ACCESSION  
 BF608664  
 VERSION  
 BF608664.1 GI:13505156  
 KEYWORDS  
 EST.  
 SOURCE  
 house mouse.  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 802)  
 Yehyawi, M., Henning, S., Neidhardt, L., Radelof, U., Hermann, B. G.,  
 Lehrich, H. and O'Brien, J.  
 TITLE  
 Detection of a high number of novel genes in a 9-day mouse embryo  
 CDNA library normalised by oligonucleotide fingerprinting  
 JOURNAL  
 Unpublished (2001)  
 COMMENT  
 Contact: Hennig S  
 Laboratory 123, dept. Lehrach  
 Max-Planck-Institut fuer Molekulare Genetik  
 Ihnestr. 63-73, D-14195 Berlin, Germany  
 Tel: +49 30 8413 1612  
 Fax: +49 30 8413 1380  
 Email: hennig@molgen.mpg.de  
 EST's are made from clones being representatives of clone clusters.  
 Clone clusters were calculated from oligonucleotide fingerprints.

FEATURES  
 source  
 1. 802  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="ICRP522B2440"  
 /clone\_lib="Mouse 9-day fetus cdna library ICRP522"  
 /tissue\_type="whole embryo"  
 /dev\_stage="embryonic 9-day"  
 /lab\_host="E. coli, XL1 blue"  
 /note="Vector: pSVSPORT1; Site\_1: NotI; Site\_2: SalI;  
 Library preparation by oligo-dt priming of RNA. Clones can  
 be ordered from the Resource Center in Berlin,  
 http://www.rzpd.de."

BASE COUNT 193 a 206 c 243 g 159 t 1 others  
 ORIGIN

Query Match 69.7%; Score 744.6; DB 12; Length 802;  
 Best Local Similarity 99.3%; Pred. No. 5.5e-109;  
 Matches 747; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 169 GGTCACTGGCCACCTGTAACCTGGCGGAGCGCGCCTCTGGAAGCCGGACAGCC 228  
 |||||||  
 Db 51 GGTCACTGGCCACCTGTAACCTGGCGGAGCGCGCCTCTGGAAGCCGGACAGCC 110  
 |||||||  
 QY 229 CCGTTTTCACAGCAGTGTGAGGCTTGGGACCCAGAAACAAAGTGAAGTCCGGC 288  
 |||||||  
 Db 111 CCGTTTTCACAGCAGTGTGAGGCTTGGGACCCAGAAACAAAGTGAAGTCCGGC 170  
 |||||||  
 QY 289 TCGTTTCCAGAGCCTGGGCGCAGCGCGCGCTGTGGAGCAGAGGTGGAGCAGCTGT 348  
 |||||||  
 Db 171 TCGTTTCCAGAGCCTGGGCGCAGCGCGCGCTGTGGAGCAGAGGTGGAGCAGCTGT 230  
 |||||||  
 QY 349 ACACCTAAAGATGAAGGCTGGGGTGGCTAGCCCTACTTTTGGGGGCTCTCTGGGAAT 408  
 |||||||  
 Db 231 ACACCTAAAGATGAAGGCTGGGGTGGCTAGCCCTACTTTTGGGGGCTCTCTGGGAAT 290  
 |||||||  
 QY 409 GCCGCGGCTGCAAGGAGCAGAGATCTACAGTGGAGCTTCCAGGCGCTCGGTGATGA 468  
 |||||||  
 Db 291 GCCGCGGCTGCAAGGAGCAGAGATCTACAGTGGAGCTTCCAGGCGCTCGGTGATGA 350  
 |||||||  
 QY 469 TTAGAGTGGAAATTTGCCCGGTGGAGCCCAAGAGACCATTCAGATGAGTCTTCCGA 528  
 |||||||  
 Db 351 TTAGAGTGGAAATTTGCCCGGTGGAGCCCAAGAGACCATTCAGATGAGTCTTCCGA 410  
 |||||||  
 QY 529 ATCAATCCAGATGGCAGCAGTCACTTGTGGAGTACTTATGCCCGCTCAGAGGCCAC 588  
 |||||||  
 Db 411 ATCAATCCAGATGGCAGCAGTCACTTGTGGAGTACTTATGCCCGCTCAGAGGCCAC 470  
 |||||||  
 QY 589 CTCACAGAGTGTCTGAGAGGAGTGTGACCAAGATGAAGAGTACGGGAGACAGATTGAC 648  
 |||||||  
 Db 471 CTCACAGAGTGTCTGAGAGGAGTGTGACCAAGATGAAGAGTACGGGAGACAGATTGAC 530  
 |||||||  
 QY 649 CCTTCTACCAACCCGCAAGATACGTACGCTGTGAGCCGGAATGGAATCCAGTGA 708  
 |||||||  
 Db 531 CCTTCTACCAACCCGCAAGATACGTACGCTGTGAGCCGGAATGGAATCCAGTGA 590  
 |||||||  
 QY 709 CTAGACTTACAGGCGCATCCGAATTGACTCAGATATCAGCGCACCCCTCAAGTTGGCTGT 768  
 |||||||  
 Db 591 CTAGACTTACAGGCGCATCCGAATTGACTCAGATATCAGCGCACCCCTCAAGTTGGCTGT 650  
 |||||||  
 QY 769 GAGAGCATTTGGAAGAAATAGAGAGTACGTTATCCGAATCTCTCCAGAGAGCTGAC 828  
 |||||||  
 Db 651 GAGAGCATTTGGAAGAAATAGAGAGTACGTTATCCGAATCTCTCCAGAGAGCTGAC 710  
 |||||||  
 QY 829 AACGTTAAAGACAACTTTCAGTAAAGGAGCAGATCATGTGACATGCCCTGCACAGA 888  
 |||||||  
 Db 711 AACGTTAAAGACAACTTTCAGTAAAGGAGCAGATCATGTGACATGCCCTGCACAGA 770  
 |||||||

QY 889 TCTCAGATGACCTATGAATCACTGGAGCAAG 920  
 DB 771 TCTCAGATGAACCTTGAATCACTGGAGCAAG 802  
 RESULT 10  
 BF535071  
 LOCUS 853 bp mRNA linear EST 11-DEC-2000  
 DEFINITION 602054331F1 NCI\_CGAP\_Sg2 Mus musculus cDNA clone IMAGE:4193478 5',  
 mRNA sequence.  
 ACCESSION BF535071  
 VERSION BF535071.1 GI:11622434  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 853)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-rt@mail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLM9525 row: n column: 07  
 High quality sequence stop: 685.  
 Location/Qualifiers  
 1..853  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone\_image="IMAGE:4193478"  
 /clone\_lib="NCI\_CGAP\_Sg2"  
 /lab\_host="DH10B (TI phage-resistant)"  
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site:1;  
 NotI; Site:2; Salt: Cloned unidirectionally. Primer: Oligo  
 dT. Average insert size 1.3 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 192 a 205 c 284 g 172 t  
 ORIGIN  
 Query Match 68.7%; Score 734; DB 12; Length 853;  
 Best Local Similarity 95.9%; Pred. No. 2.6e-107;  
 Matches 775; Conservative 0; Mismatches 30; Indels 3; Gaps 2;  
 QY 72 GAGGGCGGGCTTAAAGTCTGCTGGGGGCTCGCAGCGGACCTCAGATCTCCG 131  
 DB 1 GAGGGCGGGCTTAAAGTCTGCTGGGGGCTCGCAGCGGACCTCAGATCTCCG 60  
 QY 132 TTAGTGCTCTAAGTAAAGTCCGGGACCTGGCCAGCTGCACTGCACTGCACTG 191  
 DB 61 TTAGTGCTCTAAGTAAAGTCCGGGACCTGGCCAGCTGCACTGCACTGCACTG 120  
 QY 192 GCGGGAGCCGAGACCTCTGCGAGAACCCCGGACAGCCCGCTTTTCCAGCCAGCTGCTA 251  
 DB 121 GCGGGAGCCGAGACCTCTGCGAGAACCCCGGACAGCCCGCTTTTCCAGCCAGCTGCTA 180  
 QY 252 GGGTGGGACCCACAGAAAACAAAGTGAAGTCCGGCTGCTTTCCAGAGCTGGGCAAG 311  
 DB 181 GGGTGGGACCCACAGAAAACAAAGTGAAGTCCGGCTGCTTTCCAGAGCTGGGCAAG 240  
 QY 312 GCGGGCGCCCTGGGAGACAGAGTGAAGCACCCTGTTACACTAAGATGAAGGCTGGGG 371  
 DB 241 GCGGGCGCCCTGGGAGACAGAGTGAAGCACCCTGTTACACTAAGATGAAGGCTGGGG 300  
 QY 372 TTGGCTAGCCCTACTTTTGGGGTCTCTGTGGAACTGGCTGGGCTCGAAGGACCAAGA 431  
 TTTTGGGGTCTCTGTGGAACTGGCTGGGCTCGAAGGACCAAGA 431

DB 301 TTGGCTAGCCCTACTTTTGGGGTCTCTGTGGAACTGGCTGGGCTCGAAGGACCAAGA 360  
 QY 432 TCTACACTGTGGAGCTTGCAGGAGCTCTGGTGAATGAATGAGTGGAAATTCGCCGGT 491  
 DB 361 TCTACACTGTGGAGCTTGCAGGAGCTCTGGTGAATGAATGAGTGGAAATTCGCCGGT 420  
 QY 492 GGACCCCAAGAAAGACCATTCAGATGGGATCTTCGGAATCAATCCAGATGCGAGCAGTC 551  
 DB 421 GGACCCCAAGAAAGACCATTCAGATGGGATCTTCGGAATCAATCCAGATGCGAGCAGTC 480  
 QY 552 AGTTGTGGAGTACTTATGCCCCGCTCAGAGGCCACCTCAGAGAGTCTTGAGAGGT 611  
 DB 481 AGTTGTGGAGTACTTATGCCCCGCTCAGAGGCCACCTCAGAGAGTCTTGAGAGGT 540  
 QY 612 GTGTGACCAATGAAGAGTATGAGGAGCAAGATGACCTTACCCACCGCAAGACTA 671  
 DB 541 GTGTGACCAATGAAGAGTATGAGGAGCAAGATGACCTTACCCACCGCAAGACTA 600  
 QY 672 CGTACCGCTCTGAGCGGAGATGAGAAATCAGTGAATCACTAGACTTACAGGCGATCGAAT 731  
 DB 601 CGTACCGCTCTGAGCGGAGATGAGAAATCAGTGAATCACTAGACTTACAGGCGATCGAAT 660  
 QY 732 TGACTCAGATATCAGCGGACACCTCAAGTTGCGTGTGAGACATTTGTGAAGATACGA 791  
 DB 661 TGACTCAGATATCAGCGGACACCTCAAGTTGCGTGTGAGACATTTGTGAAGATACGA 719  
 QY 792 GATGAGCTTATGCAATTTCTCCAGAGAGGCTGACACGTTAAAGACAACTTTGCG 851  
 DB 720 GATGAGCTTATGCAATTTCTCCAGAGAGGCTGACACGTTAAAGACAACTTTGCG 777  
 QY 852 TAAGCGACAGATCTATGTGACCATGCC 879  
 DB 778 TAGGGGACAGATTTATGTGACTGCC 805  
 RESULT 11  
 B1408525  
 LOCUS 808 bp mRNA linear EST 14-AUG-2001  
 DEFINITION 60296495F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5120471 5',  
 mRNA sequence.  
 ACCESSION B1408525  
 VERSION B1408525.1 GI:15169448  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 808)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-rt@mail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: NCI\_CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLM11293 row: n column: 24  
 High quality sequence start: 4  
 High quality sequence stop: 768.  
 Location/Qualifiers  
 1..808  
 /organism="Mus musculus"  
 /strain="C2ECH 11"  
 /db\_xref="taxon:10090"  
 /clone\_image="IMAGE:5120471"  
 /clone\_lib="NCI\_CGAP\_Lu33"  
 /tissue\_type="pooled lung tumors"  
 /lab\_host="DH10B (phage-resistant)"









DB 771 CTTTATATTCGTTTATGAAATGAACTGAAAAAACTCTGAAAAACGAAAAAA 830

QY 1035 AAAAAAAAAA 1046

DB 831 AAGGAGAGAAA 842

RESULT 14

BI411532 709 bp mRNA linear EST 14-AUG-2001

LOCUS 602965126F1.NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5120574 5',

DEFINITION mRNA sequence.

VERSION BI411532.1 GI:15172455

ACCESSION BI411532.1

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 709)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L14M11294 row: c column: 07 High quality sequence start: 4 High quality sequence stop: 701.

# FEATURES

source

1. 709

Location/Qualifiers

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="IMAGE:5120574"

/clone\_1lb="NCI-CGAP\_Lu33"

/tissue\_type="pooled lung tumors"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pMT33-Pac (Pharmacia) with a modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCCGCTCTGTTTCTTTTCTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pMT33 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 191 a 170 c 214 g 132 t 2 others

ORIGIN

Query Match 64.2%; Score 686; DB 13; Length 709; Best Local Similarity 98.9%; Pred. No. 1.2e-99; Matches 700; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 256 TGGAGCCACAGAAAAAGTAGAGTCGCCGCTCTTCCAGAGCTGGCCACGGCGG 315

DB 3 TGGAGCCACAGAAAAAGTAGAGTCGCCGCTCTTCCAGAGCTGGCCACGGCGG 62

QY 316 CGGCGGTGGAGACAGAGTGGAGCACTGTACACATTAAGATGAAAGGCTGGGTTGG 375

DB 63 CGGCGGTGGAGACAGAGTGGAGCACTGTACATTAAGATGAAAGGCTGGG-7GG 121

QY 376 CTAGCCCTACTTTTGGGGTCTGCTGGAACTGCTGGGCTGGAAGGACCAAGTCTA 435

DB 122 CTAGCCCTACTTTTGGGGTCTGCTGGAACTGCTGGGCTCGGAAGCAATCTTA 181

QY 436 CACTGTGAGCTTTCAGAGGCTCTGTGTGATGATTAAGTGGGAAATTCGCCGCTGGAC 495

DB 182 CACTGTGAGCTTTCAGAGGCTCTGTGTGATGATTAAGTGGGAAATTCGCCGCTGGAC 241

QY 496 CCCAGAGACATTCATGATGGGATCTTCCGATCAATCCAGATGGACCCAGTCAATT 555

DB 242 CCCAGAGACATTCATGATGGGATCTTCCGATCAATCCAGATGGACCCAGTCAATT 301

QY 556 GTGGAGTACCTTATGAGCCGCTCAGAGGCGCCACCTCAGAGATGGTGTAGAGGTGTG 615

DB 302 GTGGAGTACCTTATGAGCCGCTCAGAGGCGCCACCTCAGAGATGGTGTAGAGGTGTG 361

QY 616 GACCGAATGAAGAGATGACGGGGAACAGATGACCTTACCCACCGCAAGAACTACGTA 675

DB 362 GACCGAATGAAGAGATGACGGGGAACAGATGACCTTACCCACCGCAAGAACTACGTA 421

QY 676 CGCGTGTAGCCGGAATGAGAAATTCAGTGAATGACTTACAGGGCATTCGAATTGAC 735

DB 422 CGCGTGTAGCCGGAATGAGAAATTCAGTGAATGACTTACAGGGCATTCGAATTGAC 481

QY 736 TCAGATATCAGCGGACACCTCAAGTTTGGCTGTGAGAGCATTTGGAAATATACGAGAT 795

DB 482 TCAGATATCAGCGGACACCTCAAGTTTGGCTGTGAGAGCATTTGGAAATATACGAGAT 541

QY 796 GAGCTTATGAAATTCCTTCACAGAGGCTGACAACTTAAGACAACTTTCAGTAAG 855

DB 542 GAGCTTATGAAATTCCTTCACAGAGGCTGACAACTTAAGACAACTTTCAGTAAG 601

QY 856 CGGACAGATCTATGTGACCATGCGCTGACAGATCTCAGATGAGTATGATCACTGGA 915

DB 602 CGGACAGATCTATGTGACCATGCGCTGACAGATCTCAGATGAGTATGATCACTGGA 661

QY 916 GCAAGCAGCTTACCAACAGTATGAGAACACCCCGAGAGGGAGGA 963

DB 662 GCAAGCAGCTTACCAACAGTATGAGAACACCCCGAGAGGGAGGA 709

RESULT 15

BQ963382 894 bp mRNA linear EST 21-AUG-2002

LOCUS AGENCOURT.10050169 NIH-MGC.134 Mus musculus cDNA clone

DEFINITION IMAGE:5508512 5', mRNA sequence.

ACCESSION BQ963382

VERSION BQ963382.1 GI:22378860

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 894)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Dr. David Rowe cDNA Library Preparation: Invitrogen Corp cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L14M14074 row: f column: 01 High quality sequence stop: 528.

# FEATURES

source

1. 894

Location/Qualifiers

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:5508512"

/clone\_1lb="NIH-MGC.134"

/tissue\_type="undifferentiated limb"

```
/lab_host="DH10B (phage-resistant)"
/Note="Vector: PCMV-SPORT6.1.cdb; Site.1: EcoRV; Site.2:
NotI; Cloned unidirectionally. Primer: Oligo dT. Average
Insert size 1.7 kb. Constructed by Reagen, Invitrogen
Corp. Note: this is a NIH MGC Library."
BASE COUNT      225 a      232 c      268 g      168 t
ORIGIN           1 others
```

```
Query Match      61.2%; Score 654.4; DB 14; Length 894;
Best Local Similarity 99.6%; Pred. No. 1e-94;
Matches 666; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 129 CGCTTAGTGCCCTAGTTAGTGGGGAACTGGCCAGCGGCTCACTGGCCACCTGAAC 188
    |||||||
Db 6 CGCTTAGTGCCCTAGTTAGTGGGGAACTGGCCAGCGGCTCACTGGCCACCTGAAC 65
QY 189 CTGGCGGGAGCCGGAGCCCTCTGAGAAAGCCGGACAGCCCGTTTTCCTCCAGCCAGCTG 248
    |||||||
Db 66 CTGGCGGGAGCCGGAGCCCTCTGAGAAAGCCGGACAGCCCGTTTTCCTCCAGCCAGCTG 125
QY 249 CTAGGGTTGGGACCAAGAAAGTGAAGTCCGGCTCTTCCAGAGCCTGGGCC 308
    |||||||
Db 126 CTAGGGTTGGGACCAAGAAAGTGAAGTCCGGCTCTTCCAGAGCCTGGGCC 185
QY 309 ACGGCGGGCCGCGTGGAGCAGAGTGGAGCACCCTGTTACACTAAAGATGAAGGCTG 368
    |||||||
Db 186 ACGGCGGGCCGCGTGGAGCAGAGTGGAGCACCCTGTTACACTAAAGATGAAGGCTG 245
QY 369 GGGTGGCTAGCCCTACTTTTGGGGCTCTCTGTGGAACTGCTGGGCTCGAAGAGCCA 428
    |||||||
Db 246 GGGTGGCTAGCCCTACTTTTGGGGCTCTCTGTGGAACTGCTGGGCTCGAAGAGCCA 305
QY 429 AGATCTACACTGTGAGGCTTCAGGGCTCTGGTGATGAATTAGAGTGGAAATTGCCG 488
    |||||||
Db 306 AGATCTACACTGTGAGGCTTCAGGGCTCTGGTGATGAATTAGAGTGGAAATTGCCG 365
QY 489 CGTGAGCCCCAAGAAGACCATTCAGATGGATCCTTCGAATCAATCAGATGCGAGCCA 548
    |||||||
Db 366 CGTGAGCCCCAAGAAGACCATTCAGATGGATCCTTCGAATCAATCAGATGCGAGCCA 425
QY 549 GTCAAGTTGTGAGGTAACCTTATGCCGCTCAGAGGCCACCTCAGAGATTGCTTGAGA 608
    |||||||
Db 426 GTCAAGTTGTGAGGTAACCTTATGCCGCTCAGAGGCCACCTCAGAGATTGCTTGAGA 485
QY 609 GGTGTGTACCGAATGAAGAGTACGGGAAACAGATTGACCTTCTACCCACCGCAAGAA 668
    |||||||
Db 486 GGTGTGTACCGAATGAAGAGTACGGGAAACAGATTGACCTTCTACCCACCGCAAGAA 545
QY 669 CTACGTACGGGTGCTGAGCCGGGAATGGAATCCAGTGAACCTTACAGGCGCATCCG 728
    |||||||
Db 546 CTACGTACGGGTGCTGAGCCGGGAATGGAATCCAGTGAACCTTACAGGCGCATCCG 605
QY 729 AATTGACTCAGATATCAGCGCACCTCAAGTTGGCTGAG -ACCATGTGGAAAGAA 787
    |||||||
Db 606 AATTGACTCAGATATCAGCGCACCTCAAGTTGGCTGAGAGCAATTGTGAAAGAA 665
QY 788 ACGAGGATG 796
    |||||||
Db 666 ACGAGGATG 674
```

Search completed: March 30, 2003, 20:56:53  
Job time : 1531.26 secs



GenCore version 5.1.4-P5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:43:09 ; Search time 1166.6 Seconds

(without alignments)  
2526.643 Million cell updates/sec

Title: US-10-082-502-19  
Perfect score: 959  
Sequence: 1 MKGGMALLLGLVLTGTA.....LCSTRDLCHALHRSDEL 182

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO.spool/US10082502/runat.24032003\_135100\_6599/app-query.fasta.1.1308  
-DB=EST -QMT=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-USER=us10082502 -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-NO.XLPHY -NO.MMAP -LARGEQUERY -NES.SCORER=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlum:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estlom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_juv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vtl:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	959	100.0	754	11 AK013568	AK013568 Mus muscu
2	959	100.0	772	11 AK007914	AK007914 Mus muscu
3	959	100.0	808	13 B1408525	B1408525 602964995
4	959	100.0	843	13 B1412539	B1412539 602990995
5	959	100.0	853	11 AK013014	AK013014 Mus muscu
6	959	100.0	1255	11 AK019927	AK019927 Mus muscu
7	947	98.7	802	12 BF608664	BF608664 MY1.00160
8	946	98.6	891	13 B1408381	B1408381 602964013
9	943	98.3	724	14 BQ210416	BQ210416 UT-R-DY1-
10	938	97.8	588	14 BM832888	BM832888 K-EST0107
11	938	97.8	591	14 BM832888	BM832888 K-EST0107
12	938	97.8	611	13 B1712899	B1712899 1d98f09.Y
13	938	97.8	697	12 BF344334	BF344334 602014708
14	938	97.8	719	12 BE733900	BE733900 601568370
15	938	97.8	742	12 BG824549	BG824549 602728454
16	938	97.8	745	12 BE791763	BE791763 601581586
17	938	97.8	1117	13 BM552907	BM552907 AGENCOURT
18	932	97.2	880	14 BQ421827	BQ421827 AGENCOURT
19	930	97.0	783	13 B1550436	B1550436 603132905
20	925	96.5	868	14 BQ421059	BQ421059 AGENCOURT
21	923	96.2	791	12 BG704443	BG704443 602688633
22	920	95.9	856	13 B1454180	B1454180 603170602
23	920	95.9	992	13 B1456021	B1456021 603170602
24	917	95.6	773	13 B1550162	B1550162 603193045
25	913	95.2	709	13 B1411532	B1411532 602965126
26	912	95.1	878	13 B1763950	B1763950 603049829
27	913	95.1	810	12 BE870315	BE870315 601447464
28	911	95.0	887	13 B1762759	B1762759 603048045
29	910	94.9	665	10 AW919569	AW919569 EST350873
30	906	94.5	720	13 B1909759	B1909759 603070841
31	905	94.4	673	14 BQ668530	BQ668530 AGENCOURT
32	898	93.6	783	13 B1763784	B1763784 603049729
33	898	93.6	914	12 BE799445	BE799445 601889804
34	891	92.9	765	12 BF312281	BF312281 601889804
35	886	92.4	678	10 BE309953	BE309953 601091903
36	884	92.2	856	12 BF965088	BF965088 602547535
37	878	91.6	751	12 BE742698	BE742698 601574753
38	874	91.1	657	14 BQ211867	BQ211867 UT-R-DY1-
39	874	91.1	945	13 B1412341	B1412341 602991837
40	870	90.7	867	13 BG973691	BG973691 602844540
41	870	90.7	885	14 BQ920767	BQ920767 AGENCOURT
42	870	90.7	963	14 BQ929020	BQ929020 AGENCOURT
43	869	90.6	674	14 BQ202181	BQ202181 UT-R-EB1-
44	868	90.5	797	13 B1762927	B1762927 603047987
45	866	90.3	549	10 BE198865	BE198865 u983d11.Y

## ALIGNMENTS

RESULT 1  
AK013568  
LOCUS  
DEFINITION  
AK013568 754 bp mRNA linear HTC 19-JAN-2002  
Mus musculus adult male hippocampus cDNA, RIKEN full-length  
enriched library, clone:2900022B12;transmembrane protein 4, full  
insert sequence.  
ACCESSION  
AK013568  
VERSION  
AK013568.1 GI:12850979  
KEYWORDS  
HTC; CAP trapper  
SOURCE  
Mus musculus (strain:C57BL/6J) adult male hippocampus cDNA to mRNA,  
clone:2900022B12.  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.





transmembrane protein 4"  
 /codon\_start=1  
 /protein\_id="BAB25346.1"  
 /db\_xref="GI:12841774"  
 /db\_xref="MGI:1928477"  
 /translation="MKMGWLLLLGLVLTAMARRSODLHGCACRAVDLEMETAR  
 VDPKRTIOMGSRFRINPDGQSVPEVPYPASEAHLRELEVCDDRMKEVEQIDPSTHR  
 KATRVASRNGSSSELDQGIINDSIDISGLTFACBSLYEEDELEIEFSREADNVK  
 DKLSKRIDLCDHALHRSDEL"

BASE COUNT 210 a 180 c 225 g 157 t

ORIGIN

Alignment Scores:  
 Pred. No.: 5.55e-108 Length: 772  
 Score: 959.00 Matches: 182  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 11 Gaps: 0

US-10-082-502-19 (1-182) x AK007914 (1-772)

QY 1 Metlysglytrpqltytrpleualaleuleuenglyvalleuleuglythraltatpala 20  
 |||||  
 Db 104 ATGAAGGCTGGGGTGGCTAGCCCTACTTTGGGGGCTCCTGGGAAGCTGCTGGGCT 163  
 QY 21 ArgArgSerGlnAspLeuHscysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 |||||  
 Db 164 CGAAGAGCCAGATCTACACTGTGAGGCTTGCAAGGCTCTGTGTGATGAATTAAGATGG 223  
 QY 41 GluIleAlaArgValAspProlylsThrIleGlnMetClySerPheArgIleAsnPro 60  
 |||||  
 Db 224 GAATATGCCCGCTGGAGCCCAAGAAAGACCATTCAGATGGAGATCCCTCCAAATCAATCCA 283  
 QY 61 AspGlySerGlnSerValAlaGluValProtyrAlaArgSerGluAlaHisLeuThrGlu 80  
 |||||  
 Db 284 GATGGACGACGACGACGATGTGGAGTACCTTATGCCCGCTCAAGGCCCACTCAGACAG 343  
 QY 81 LeuLeuGluGluValAlaCysAspArgMetlysglytrtyrGlyGluGlnIleAspProSerThr 100  
 |||||  
 Db 344 TTGCTTGAGAGAGGTGTGTGACCGAATGAAGAGTACGAGGAGACAGATTCACCTTCTTACC 403  
 QY 101 HisArgLysAsnTrpValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 120  
 |||||  
 Db 404 CACCGCAAGACTACGCTAGCGCTGAGCCGGAATGGAGAAATCCAGTGAATAGACTTA 463  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 |||||  
 Db 464 CAGGCGATCCGATTCAGATACAGATATCAGGCGACCCCTCAAGTTTGCCTGTGAGACATT 523  
 QY 141 ValGluGluTrpGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
 |||||  
 Db 524 GTGGGAAGATACAGAGATGAGCTTATCGAATTCCTTCCAGAGAGGCTGCAACAGCTTAA 583  
 QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 |||||  
 Db 584 GAAACACTTGGCAGTAAGGAGACAGATCTATGTGACATGCTCGACAGATCTCAGAT 643  
 QY 181 GluLeu 182  
 |||||  
 Db 644 GAGCTA 649

RESULT 3  
 BI408525 808 bp mRNA linear EST 14-AUG-2001  
 LOCUS 60296495F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5120471 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI408525  
 VERSION BI408525.1 GI:15169448  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
 1 (bases 1 to 808)  
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.

CDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM1293 row: n column: 24  
 High quality sequence start: 4  
 High quality sequence stop: 768.  
 Location/Qualifiers

FEATURES  
 source  
 1..808  
 /organism="Mus musculus"  
 /strain="CZECH 11"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5120471"  
 /clone\_1lb="NCI\_CGAP\_Lu33"  
 /tissue\_type="pooled lung tumors"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: lung; Vector: pTZ19-Pac (Pharmacia) with a  
 modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st  
 strand cDNA was prepared from mRNA obtained from pooled  
 lung tumors with a Not I - oligo(dT) primer [5'  
 TGTACCAATCTGAAGTGGAGCGGCGCTCTGTTTCTTTTCTTTT 3'].  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not  
 I and Eco RI sites of the modified pTZ19 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 224 a 181 c 237 g 166 t

ORIGIN

Alignment Scores:  
 Pred. No.: 5.93e-108 Length: 808  
 Score: 959.00 Matches: 182  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-10-082-502-19 (1-182) x BI408525 (1-808)

QY 1 Metlysglytrpqltytrpleualaleuleuenglyvalleuleuglythraltatpala 20  
 |||||  
 Db 119 ATGAAGGCTGGGGTGGCTAGCCCTACTTTGGGGGCTCCTGGGAAGCTGCTGGGCT 178  
 QY 21 ArgArgSerGlnAspLeuHscysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 |||||  
 Db 179 CGAAGAGCCAGATCTACACTGTGAGGCTTGCAAGGCTCTGTGTGATGAATTAAGATGG 238  
 QY 41 GluIleAlaArgValAspProlylsThrIleGlnMetClySerPheArgIleAsnPro 60  
 |||||  
 Db 239 GAATATGCCCGCTGGAGCCCAAGAAAGACCATTCAGATGGAGATCCCTCCAAATCAATCCA 298  
 QY 61 AspGlySerGlnSerValAlaGluValProtyrAlaArgSerGluAlaHisLeuThrGlu 80  
 |||||  
 Db 299 GATGGACGACGACGACGATGTGGAGTACCTTATGCCCGCTCAAGGCCCACTCAGACAG 358  
 QY 81 LeuLeuGluGluValAlaCysAspArgMetlysglytrtyrGlyGluGlnIleAspProSerThr 100  
 |||||  
 Db 359 TTGCTTGAGAGAGGTGTGTGACCGAATGAAGAGTACGAGGAGACAGATTCAGATCTTACC 418  
 QY 101 HisArgLysAsnTrpValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 120  
 |||||  
 Db 419 CACCGCAAGACTACGCTAGCGCTGAGCCGGAATGGAGAAATCCAGTGAATAGACTTA 478



QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 479 CAGGCGATCCGAAATGACTACATATCAGCGACCCCTCAAGTTTCGCTGAGACATT 538  
 QY 141 ValGluGluThrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
 DB 539 GTGGAAGAATACAGAGATAGCTTATCGAATCTTCTCCAGACAGGCTGACAACTTTAA 598  
 QY 161 AspValLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 DB 599 GACAAACTTTCGAGTAAGCGGACAGATCTATGTGACCATGCCCTGCACAGATCTCAGAT 658  
 QY 181 GluLeu 182  
 DB 659 GAACCT 664  
 RESULT 4  
 LOCUS B1412539 843 bp mRNA linear EST 14-AUG-2001  
 DEFINITION 60290995F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5146967 5',  
 mRNA sequence.  
 ACCESSION B1412539  
 VERSION B1412539.1 GI:15173462  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 NIH-MGC http://mgc.nci.nih.gov/  
 1 (bases 1 to 843)  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaab@femail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 place: LMAH1362 row: n column: 24  
 High quality sequence start: 38  
 High quality sequence stop: 824.  
 Location/Qualifiers  
 1..843  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5146967"  
 /clone\_1lb="NCI\_CGAP\_Lu33"  
 /tissue\_type="pooled lung tumors"  
 /lab\_host="pH10B (phage-resistant)"  
 /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: NotI; Site\_2: EcoRI; 1st  
 strand cDNA was prepared from mRNA obtained from pooled  
 lung tumors with a Not I - oligo(dT) primer [5'  
 TGTTCACATCTGAGTGGAGCGGCGCCCTGCTGTTTCTTTTCTTTT 3']  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not  
 I and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo. "  
 BASE COUNT 239 a 195 c 238 g 171 t  
 ORIGIN  
 Alignment Scores:  
 Pctd. No.: 6.3e-108 Length: 843  
 Score: 959.00 Matches: 182  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0  
 US-10-082-502-19 (1-182) x B1412539 (1-843)  
 QY 1 MetLysGlyThrPglYThrPleuAlaLeuLeuLysGlyValLeuLysGlyThrAlaTrpAla 20  
 DB 154 ATGAAAGGCTGGGGTGTGGCTAGCCCTACCTTTTGGGGGCTCTCGGAGACTCGGGCT 213  
 QY 21 ArgArgSerGluAspLeuHisArgGlyValAcysArgAlaLeuValAspAluLeuGluTrp 40  
 DB 214 CGAAGAGCGCAAGATCTACACTGTGGAGCTTGCAGGCGCTGTGGATAAATTAGAGTGG 273  
 QY 41 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 274 GAATATGCCCGGGTGGAGCCCAAGAGACCATTCAGATGGATCTTCGATCAATCA 333  
 QY 61 AspGlySerGlnSerValValGluValProTyralaArgSerGluAlaHisLeuThrGlu 80  
 DB 334 GATGGCAGCCAGTCACTTGTGGAGTACCTTATGCCCCCTCAGAGGCCACCTCAGAG 393  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 100  
 DB 394 TTGCTTGAGGAGGTGTGTACCGAATGAAGAGTACGGGGAACAGATTACCTTCTTACC 453  
 QY 101 HisArgLysAsnTyrrValArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
 DB 454 CACCGCAAGACACTAGCTAGCGCTGAGCCGGAATGAGAGATCCAGTAACTAGACTTA 513  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 514 CAGGCGATCCGAAATGACTACATATCAGCGACCCCTCAAGTTTCCGTGAGACATT 573  
 QY 141 ValGluGluThrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
 DB 574 GTGGAAGAATACAGAGATAGCTTATCGAATCTTCTCCAGAGAGGCTCACAACGTTAA 633  
 QY 161 AspValLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 DB 634 GACAAACTTTCGAGTAAGCGGACAGATCTATGTGACCATGCCCTGCACAGATCTCAGAT 693  
 QY 181 GluLeu 182  
 DB 694 GAACCTA 699  
 RESULT 5  
 LOCUS AK013014 853 bp mRNA linear HTC 19-JAN-2002  
 DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length  
 enriched library, clone:2810406L24:transmembrane protein 4, full  
 insert sequence.  
 ACCESSION AK013014  
 VERSION AK013014.1 GI:12850114  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (strain:C57BL/6J) 10, 11 days embryo cDNA to mRNA,  
 clone:1lb:RIKEN full-length enriched mouse cDNA library  
 clone:2810406L24.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 REFERENCES  
 1 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 2049374



DB	634	GENGAAGATACGAGGATGAGCTTATCCGAATTCCTCCACGAGAGCGAGACAGCTTAA	693
OY	161	Asp1y5IeucY5e5e1y5a7y3rAsp1eucY5a5pH15a1a1euh15a75e1rH15a1p	180
DB	694	GACAAACCTTTCAGTAAACGGACAGATCTATGTGACCATGCGCCCTGCACAGATCTCAGAT	753
OY	181	Glu1euc	182
DB	754	GAGCTA	759
RESULT 6			
LOCUS	AK019927	1255 bp	mRNA
DEFINITION	Mus musculus adult male pituitary gland cDNA, RIKEN full-length enriched library, clone:5330432A10:transmembrane protein 4, full insert sequence.		HTC 19-JAN-2002
ACCESSION	AK019927		
VERSION	AK019927.1	GI:12860332	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) adult male pituitary gland cDNA to mRNA, clone:lib.RIKEN full-length enriched mouse cDNA library clone:5330432A10.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS			
TITLE			
JOURNAL	Carninci, P. and Hayashizaki, Y.		
MEDLINE	High-efficiency full-length cDNA cloning		
PUBMED	Meth. Enzymol. 303, 19-44 (1999)		
REFERENCE	99279253		
AUTHORS	10349636		
TITLE			
JOURNAL	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
MEDLINE	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.		
PUBMED	Normalization and subtraction of cap-trapper-selected cDNAs to		
REFERENCE	prepare full-length cDNA libraries for rapid discovery of new genes		
AUTHORS	Genome Res. 10 (10), 1617-1630 (2000)		
TITLE			
JOURNAL	20499374		
MEDLINE	11042159		
PUBMED	3		
REFERENCE			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		
TITLE	Sun, N. H., Akiyama, J., Nishi, K., Kitanishi, T., Tashiro, H., Itoh, M.,		
JOURNAL	Komi, H., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,		
MEDLINE	Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,		
PUBMED	Fujieda, S., Inoue, K., Togawa, Y., Iizawa, M., Ohara, E., Watahiki, M.,		
REFERENCE	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,		
AUTHORS	Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format		
JOURNAL	sequencing pipeline with 384 multiplexillary sequencer		
MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)		
PUBMED	11076861		
REFERENCE	4		
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,		
TITLE	Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,		
JOURNAL	Alkawa, K., Iizawa, M., Nishi, K., Kiyosawa, H., Kondou, S., Yamanaka, I.,		
MEDLINE	Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R.,		
PUBMED	Kadoi, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,		
REFERENCE	Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochwa, H.,		
AUTHORS	Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,		
TITLE	Queckenbush, J., Schiml, L. M., Stäubli, F., Suzuki, R., Tomita, M.,		
JOURNAL	Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,		
MEDLINE	Baderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,		
PUBMED	Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bull, C.,		
REFERENCE	Fletcher, C., Fujita, M., Garbolino, M., Gustincich, S., Hill, D.,		
AUTHORS	Hofmann, M., Hume, D. A., Kamuya, M., Lee, N. H., Lyons, P.,		
TITLE	Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,		
JOURNAL	Rling, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,		
MEDLINE	Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K. F., Suzuki, H.,		
PUBMED	Toyo-Oka, K., Wang, K. H., Weitz, C., Willeker, C., Wilmung, L.,		
REFERENCE	Wyrshav-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S.		
AUTHORS	and Hayashizaki, Y.		
TITLE	Functional annotation of a full-length mouse cDNA collection		

JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085560
PUBMED	1121/7851
AUTHORS	5 (bases 1 to 1255) Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arawaka,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Cantinci,P., Fukuda,S., Fukuishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hironaka,T., Horii,F., Hume,D., Imobanti,K., Ishii,Y., Itoh,H., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kushihara,C., Matsuyama,T., Miyazaki,A., Nishii,K., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Saeki,D., Schirml,L., Shidata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toyota,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute: 1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp). URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216) Please visit our web site ( <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> ) for further details.
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAAGACATCCAGACGCCTCTTTTCTTTTTTCTTCA 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trap. cDNA went through one round of normalization to Rot = 3.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adaptor of sequence [5' GAGAGAAGATTCGACGTAAATTAAATTAATCATCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBlueScript KS(+) after bulk excision from lambda Phi C I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.
FEATURES	Location/Qualifiers
SOURCE	1..1255 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="PANTO:DB:5330432A10" /db_xref="MGD:MGI:1899093" /db_xref="taxon:10090" /clone="5330432A10" /sex="male" /tissue_type="pituitary gland" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 1..1255 /gene="Rmem4" 215..763 /gene="Rmem4" /note="data source:MGD, source key:MGI:1928477, evidence:ISS putative transmembrane protein 4" /codon_start=1 /protein_id="BAB31921.1" /db_xref="GI:12860333" /db_xref="MGD:MGI:1928477" /translation="MKGMGLALLGLVLTGAARRSODLHCAGCRALVDLEWELAR VDPKRTIOMGSERINPDGSSVVEVPARSANHTLEILEFCMDKRGQIDPSSTRH KNRYRRSRNGSESLDIQGRIDSISGLTFACEISIVEYEDELIEFESREADNVK DKLSKRDTDLDAHLRSHEL"
GENE	
CDS	
BASE COUNT	309 a 293 c 354 g 299 t
ORIGIN	

## Alignment Scores:

Pred. No.: 1.12e-107 Length: 1255  
 Score: 959.00 Matches: 182  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 11 Gaps: 0

US-10-082-502-19 (1-182) x AK019927 (1-1255)

QY 1 MettysGlyTrrpGlyTrrpLeuLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
 DB 215 ATGAAAGGCTGGGCTGGCTAGCCCTACTTTGGGGCTCTGCGGAAGCTGCTGGGCT 274  
 QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 DB 275 CGAAGAGCCAAAGATCTACACTGTGAGCTTGAGGCGCTGTGTGATGAATTAGAGTGG 334  
 QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 60  
 DB 335 GAATTTGCCCGGCTGGAGCCCAAGAACACATTCAGATGGATCCCTCCGAATCAATCCA 394  
 QY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80  
 DB 395 GATGGCAGCCAGCTCACTTGTGAGGTACCTTATGCCGCTCAGAGGCCACCTCAGACAG 454  
 QY 81 LeuLeuGluGluValAlaCysAspArgMetLysGlyTyrGlyGlnIleAspProSerThr 100  
 DB 455 TTGCTTGAGGAGGTGTGTGACCAATGAAGAGATACGGGAAACAGATTGACCTTCTAC 514  
 QY 101 HisArgGlyAsnTyrValArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
 DB 515 CACCGCAAGACTACGTACGCTGCGAGCCGGAAGAGAAATCCAGTACTAGACTTA 574  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 575 CAGGCGATCCGAATTCACATACATACAGCGCACCCCTCAAGTTTCGTGTGAGACATT 634  
 QY 141 ValGluGluTyrGluAspGluLeuIleGlnPheSerArgGluAlaAspAsnValLys 160  
 DB 635 GTGGAAAGATACGAGGATGAGCTTATCGAATCTCTCCAGAGAGCTGACAACGTTAA 694  
 QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 DB 695 GACAAACTTGGAGTAAAGGAGACAGATCTATGTGACCAATGCCCTGACAGATCTCAGAT 754  
 QY 181 GluLeu 182  
 DB 755 GAGCTA 760

RESULT 7  
 BF608664 802 bp mRNA linear EST 01-APR-2001  
 LOCUS BF608664 MYL1001606 Mouse 9-day fetus cDNA library ICRF522 Mus musculus  
 DEFINITION cDNA clone ICRF522B2440 5', mRNA sequence.  
 ACCESSION BF608664  
 VERSION BF608664.1 GI:13505156  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 802)  
 AUTHORS Yehyawi, M., Hennig, S., Neldhardt, L., Radelof, U., Hermann, B. G.,  
 Lehrach, H. and O'Brien, J.  
 TITLE Detection of a high number of novel genes in a 9-day mouse embryo  
 cDNA library normalised by oligonucleotide fingerprinting  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Hennig S  
 Laboratory 123, dept. Lehrach  
 Max-Planck-Institut fuer Molekulare Genetik  
 Ihnestr. 63-73, D-14195 Berlin, Germany

## FEATURES

## source

Tel: +49 30 8413 1612  
 Fax: +49 30 8413 1380  
 Email: hennig@molgen.mpg.de  
 EST's are made from clones being representatives of clone clusters.  
 Clone clusters were calculated from oligonucleotide fingerprints.  
 PCR primers  
 FORWARD: 5'-GAGCTATTCAGACAGTACTGCA-3'  
 BACKWARD: 5'-TAATACGACTCTACATAGAG-3'  
 Seq primer: 5'-ATTAGTGACACTATAG-3'  
 High quality sequence stop: 802.  
 Location/Qualifiers  
 1..802  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="ICRF522B2440"  
 /clone\_1ib="Mouse 9-day fetus cDNA library ICRF522"  
 /tissue\_type="whole embryo"  
 /dev\_stage="embryonic 9-day"  
 /lab\_host="E.coli, XL1 blue"  
 /note="Vector: pSVSPORT1; Site\_1: NotI; Site\_2: SalI;  
 library preparation by oligo-dT priming of RNA. Clones can  
 be ordered from the Resource Center in Berlin,  
 http://www.rzpd.de."

BASE COUNT 193 a 206 c 243 g 159 t 1 others  
 ORIGIN

## Alignment Scores:

Pred. No.: 1.78e-106 Length: 802  
 Score: 947.00 Matches: 180  
 Percent Similarity: 99.45% Conservative: 1  
 Best Local Similarity: 98.90% Mismatches: 1  
 Query Match: 98.75% Indels: 0  
 DB: 12 Gaps: 0

US-10-082-502-19 (1-182) x BF608664 (1-802)

QY 1 MettysGlyTrrpGlyTrrpLeuLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
 DB 240 ATGAAAGGCTGGGCTGGCTAGCCCTACTTTGGGGCTCTGCGGAAGCTGCTGGGCT 299  
 QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 DB 300 CGAAGAGCCAAAGATCTACACTGTGAGCTTGAGGCGCTGTGTGATGAATTAGAGTGG 359  
 QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 60  
 DB 360 GAATTTGCCCGGCTGGAGCCCAAGAACACATTCAGATGGATCCCTCCGAATCAATCCA 419  
 QY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80  
 DB 420 GATGGCAGCCAGCTCACTTGTGAGGTACCTTATGCCGCTCAGAGGCCACCTCAGACAG 479  
 QY 81 LeuLeuGluGluValAlaCysAspArgMetLysGlyTyrGlyGlnIleAspProSerThr 100  
 DB 480 TTGCTTGAGGAGGTGTGTGACCAATGAAGAGATACGGTGAACAGATTGACCTTCTAC 539  
 QY 101 HisArgGlyAsnTyrValArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
 DB 540 CACCGCAAGACTACGTACGCTGAGCGCGGAATGAGAACTGCACTAGACTTA 599  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 600 CAGGCGATCCGAATTCACATACATACAGCGCACCCCTCAAGTTTCGTGTGAGACATT 659  
 QY 141 ValGluGluTyrGluAspGluLeuIleGlnPheSerArgGluAlaAspAsnValLys 160  
 DB 660 GTGGAAAGATACGAGGATGAGCTTATCGAATCTCTCCAGAGAGCTGAAACGTTAA 719  
 QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 DB 720 GACAAACTTGGAGTAAAGGAGACAGATCTATGTGACCAATGCCCTGACAGATCTCAGAT 779  
 QY 181 GluLeu 182

```

Db      780 GAACTT 785
RESULT 8
BI408381
LOCUS
DEFINITION      891 bp  mRNA  linear  EST 14-AUG-2001
602964013F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5119563 5',
mRNA sequence.
ACCESSION      BI408381
VERSION        BI408381.1  GI:15169304
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Email: Robert Strausberg, Ph.D.
Contact: Robert Strausberg, Ph.D.
Email: c9apds-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: LLM11291 row: 1 column: 04
High quality sequence stop: 873.
FEATURES
source
1.891
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_1fb="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DHI10B (phage-resistant)"
/Note="Organ: Lung; Vector: pTV3D-Pac (Pharmacia) with a
modified polylinker; Site: 1: NotI; Site 2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - Oligo(dT) primer [5'
TCCTACCACTTCGAACTGCGAGCGCCCTCTCTTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pTV3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      240 a 215 c 258 g 178 t
ORIGIN
Alignment Scores:
Pred. No.:      2,76e-106      Length:      891
Score:          946.00         Matches:      182
Percent Similarity: 99.45%      Conservative: 0
Best Local Similarity: 99.45%      Mismatches: 0
Query Match:    98.64%         Indels:      1
DB:             13            Gaps:        0
US-10-082-502-19 (1-182) x BI408381 (1-891)
QY      1 MellySGlyTTPGlyTTPLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTTPAla 20
Db      198 ATGAAAGCGTGGGCTGGCTAGCCCTACTTTTGGGGCTCTGCGGAACTGCGTGGCT 257
QY      21 ArgArgSerGlnAspLeuHiscysGlyAlaCysArgAlaLeuValAspGluLeuGluTTP 40
Db      258 CGAAGAGCCCAAGATCTACACTGTGAGCTTGCAAGGCTCTGGGTGAATTAAGAGCTGG 317
QY      41 GlnTlAlaAlaArgValAspProLysThrTlGlnMetGlySerPheArgIleAsnPro 60

```

```

Db      318 GAAATGCCCCCGCTGGACCCCAAGAACCATTTAGATGGGATCTTCGGAATCAATCA 377
QY      61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
Db      378 GATGACAGCCAGCTCAGTTGTGGAGTACCTTATCCCGCTCAGAGGCCCACTCAGAG 437
QY      81 LeuLeuGluGluValAlaCysAspArgMetGlyGluTyrGlyGluGlnIleAsnProSerThr 100
Db      438 TTGCTTGAAGAGGTGTGTGACCGAATGAAGAGTACGGGAAACAGATTGACCTTCAC 497
QY      101 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 120
Db      498 CACCGCAAGAACCTACAGTACGCGCTGTGATCCGGAATGAGAAATCAATGAATCAATGACTTA 557
QY      121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAla-CysGluSerIle 140
Db      558 CAGGCGAATCCGAATGACTCGATATCAGCGCACCTCAAGTTGGCTGTGAGACAT 617
QY      140 eValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValIly 160
Db      618 TGTGGAAGAAATACAGAGATGAGCTTATCGAATCTCTCCAGAGAGCTGCACACGTTAA 677
QY      160 SASpLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAs 180
Db      678 AGACAAACTTGGCAGTAAGCGACAGATGTATGACCATGCCCTGCACAGATCTCAGA 737
QY      180 pGluLeu 182
Db      738 TGAACCTT 744
RESULT 9
BQ210416/c 724 bp  mRNA  linear  EST 02-MAY-2002
DEFINITION      UI-R-DY1-col-e-06-0-UI-s1 UI-R-DY1 Rattus norvegicus cDNA clone
BQ210416
UI-R-DY1-col-e-06-0-UI 3', mRNA sequence.
VERSION        BQ210416.1  GI:20426861
KEYWORDS
SOURCE
ORGANISM
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 724)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
COMMENT
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
normalized cartilage library cDNA library preparation: M.B. Soares
Lab Clone distribution: clones will be available through Research
Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=yes.
FEATURES
source
1.724
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DY1-col-e-06-0-UI"
/clone_1fb="UI-R-DY1"

```

/tissue\_type="Cartilage"  
 /dev\_stage="ADULT"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Organ: Femur and Tibia; Vector: pRT3D-Pac  
 (Pharmacia) with a modified polylinker; Site 1: Not I;  
 Site 2: Eco RI; UT-R-DYI is a normalized cDNA library  
 containing the following tissue(s): Rat Cartilage from  
 Femur and Tibia. The library was constructed according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. First strand cDNA synthesis was primed with an  
 oligo-dT primer containing a Not I site. Double stranded  
 cDNA was ligated to an EcoR I adaptor, digested with Not  
 I, and cloned directionally into pRT73-Pac vector. The  
 oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tag for this library is CTAAATGACG. The Rat  
 cartilage tissue was provided by Dr Jeff Stevens at the  
 University of Iowa.  
 TAG\_Lib-UI-R-DYI  
 TAG\_TISSUE-cartilage  
 TAG\_SEQ-CTAAATGACG"

BASE COUNT 156 a 206 c 157 g 205 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.79e-106 Length: 724  
 Score: 943.00 Matches: 178  
 Percent Similarity: 98.35% Conservative: 1  
 Best Local Similarity: 97.80% Mismatches: 3  
 Query Match: 98.33% Indels: 0  
 DB: 14 Gaps: 0

US-10-082-502-19 (1-182) x BQ210416 (1-724)  
 Oy 1 Mettysgltgtrglttrpleuallaleuenglythralatpala 20  
 Db 669 ATGAAAGGCTGGGGTGGCTGGCTTACTTTGGGGGTCTCTGGAGACTACCTGGGCT 630  
 Oy 21 ATGATGSeRglnAspLeuNHsCysglYlaCysARgAlaLeuValAspJuleuNlurp 40  
 Db 629 CGAAGAGACGAGACTACACTGTGAGCTTCAGGGCTCTGGATGAATAGAGTGG 570  
 Oy 41 GlutleAlaAlaValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 Db 569 GAAATGGCCGGGTGGACCCCAAGAACCAATTCAGATGGATCCCTCCGAATCAATCCA 510  
 Oy 61 AspGlySerGlnSerValAlaGluValProLYrAlaArgSerGluAlaHisLeuThrglu 80  
 Db 509 GATGGCAGCCAGTCACTGTGTGAGGTCTTATGCCCTCAGAGGCCACCTCAGGAG 450  
 Oy 81 LeuLeuGluGluValCysAspArgMetLysGluTYrGlyGluGlnIleAspProSerThr 100  
 Db 449 TTGCTTGAGAGATATGTGACCGCATGAAGAGATGGGAGAACAGATTGACCTTCTACC 390  
 Oy 101 HisArgLysAsnTYrValAlaYrAlaValSerArgAsnGlyLysSerGluLeuAspLeu 120  
 Db 369 CACCCCAAGAACTACGTCGCTGTGTCGGCCGACGAGGAGATCCAGTCACTGACTTA 330  
 Oy 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 Db 339 CAGGGTATCCGAAATCGATTCAAGATTCAGTGGCACCCTCAACTTGGCTGTAGAGCAATT 270  
 Oy 141 ValGluGluTYrGluAspJuleuNlurpPheSerArgGluAlaAspAsnValLys 160  
 Db 269 GTGAGAGAGTACGAGAGTAACTTATTAATCTTTCAAGAGAGCGACAGACTTAA 210  
 Oy 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisasp 180  
 Db 209 GACCAAACTTTGAGTAAACGACAGATCTATGTGACATGCCCTCAGAGATCTCATGAT 150  
 Oy 181 GluLeu 182  
 |||||

Db 149 GAGCTA 144  
 RESULT 10  
 LOCUS BM832828  
 DEFINITION K-EST0107375 S55NU484s1 Homo sapiens cDNA clone S55NU484s1-11-B03  
 ACCESSION BM832828  
 VERSION BM832828.1 GI:19189237  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 588)  
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribd.re.kr  
 Plate: 11 row: B column: 03  
 High quality sequence stop: 588.  
 FEATURES  
 SOURCE  
 Location/Qualifiers  
 1..588  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="S55NU484s1-11-B03"  
 /clone\_lib="S55NU484s1"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_line="Epi1hel1a1"  
 /cell\_line="SNU-484"  
 /lab\_host="Top10F"  
 /note="Organ: Stomach; Vector: pTZ19BP1; Site1: EcoRI;  
 Site2: NotI. The poly (A) + RNA was decapped with tabacco  
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
 including EcoRI site by treatment of T4 DNA ligase. The  
 first strand cDNA was synthesized from oligo dt-selected  
 mRNA by priming with dt-tailed vector. The dt-tailed  
 vector was adjusted to have about 60nt. The cDNA vector  
 was circularized with E. coli DNA ligase after digestion  
 of EcoRI which site is also included in vector. An RNA  
 strand converted to a DNA strand by Okayama-Berg method.  
 The obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10F by electroporation method.  
 After analyzing and sequencing about 2,000 - 3,000  
 colonies in original cDNA library, the abundant cDNAs were  
 selected and amplified by PCR reaction using vector region  
 primer including T7 promoter as 5' primer and N(dT)14 as  
 3' primer. The PCR products were used as template for  
 synthesis of biotinylated single stranded RNA by in vitro  
 transcription reaction. The synthesized RNA probes were  
 hybridized with antisense single stranded cDNAs prepared  
 from original library and incubated with avidin-gel.  
 After removing DNA-RNA hybrids by centrifuge, the  
 subtracted cDNA libraries were constructed by  
 transformation of the remaining DNA into competent cells E.  
 coli Top10F with electroporation method."

BASE COUNT 149 a 140 c 176 g 123 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.47e-105 Length: 588  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3

Query Match: 97.81% Indels: 0  
 DB: 14 Gaps: 0  
 US-10-082-502-19 (1-182) x BM832888 (1-588)

QY 1 MetLysGlyTyrPglYTrpLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
 |||||  
 DB 2 ATGAAGAAGGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 61  
 QY 21 ArgArgSerGlnAspLeuHscysGlyAlaCysArgAlaLeuValAspLeuLeuGlyTrp 40  
 |||||  
 DB 62 CGAGAGAGCGAGAGATCCACTGTGGAGCATGACAGGCGCTGTGTGATCACTAGAAATGG 121  
 QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 |||||  
 DB 122 GAAATTTGCCAGGTGACCCCAAGAACACATTCAGATGGATCTTCCGATCAATCA 181  
 QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
 |||||  
 DB 182 GATGGCAGCCACTGATGGTGGAGTGGCTTATGCCCTCAGAGGCCACCTCCACAGAG 241  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100  
 |||||  
 DB 242 CTGCTGGAGGAGATATGTGACCGATGAAAGGATGTGGGAAACAGATTGATCCTCCACC 301  
 QY 101 HisArgLysAsnTyrValArgValValSerArgAsnGlyGlySerGluLeuAspLeu 120  
 |||||  
 DB 302 CATGCCAAGAACTACGTACGTACTGTAGTGGCGCGAATGAGAAATCCAGTGAACCTGACCTA 361  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 |||||  
 DB 362 CAAAGCAATCCCAATCGACTCAGATTTAGCGGACACCCCTCAAGTTTGGCTGTGACAGCTT 421  
 QY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
 |||||  
 DB 422 GTGAGAGAAATCGAGAGATGAATCAATTTCTTCCAGAGCGCTGACAAATGTTAAA 481  
 QY 161 AspLysLeuGlySerLysArgThrAspLeuGlyAspHisAlaLeuHisArgSerHisAsp 180  
 |||||  
 DB 482 GACAAACTTTCCACTGAAGCGAAGACATCTTGTGACCATGCCCTGCACATATCCCATGAT 541  
 QY 181 GluLeu 182  
 |||||  
 DB 542 GAGCTA 547

RESULT 11  
 LOCUS BM832888 591 bp mRNA linear EST 06-MAR-2002  
 DEFINITION K-EST0107445 SSSNU484s1 Homo sapiens cDNA clone SSSNU484s1-11-H01  
 5', mRNA sequence.  
 ACCESSION BM832888  
 VERSION BM832888.1 GI:19189297  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 591)  
 Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
 Kim, Y.S.  
 TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsungemall.kribb.re.kr  
 Plate: 11 row: H column: 01  
 High quality sequence stop: 591.  
 FEATURES Location/Qualifiers

source 1..591  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="SSSNU484s1-11-H01"  
 /clone\_1id="SSSNU484s1"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_type="Epithelial"  
 /lab\_host="SNU-484"  
 /lab\_host="TOP10F"  
 /note="Organ: Stomach; Vector: pTZ19mp1; site\_1: EcoRI;  
 site\_2: NotI; The poly (A) + RNA was decapped with tobacco  
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
 including EcoRI site by treatment of T4 DNA ligase. The  
 first strand cDNA was synthesized from oligo dt-selected  
 mRNA by priming with dt-tailed vector. The dt-tailed  
 vector was adjusted to have about 60nt. The cDNA vector  
 was circularized with E. coli DNA ligase after digestion  
 of EcoRI which site is also included in vector. An RNA  
 strand converted to a DNA strand by Okazaki-Berg method.  
 The obtained cDNA vectors were used for transformation of  
 competent cells E. coli TOP10F by electroporation method.  
 After analyzing and sequencing about 2,000 - 3,000  
 colonies in original cDNA library, the abundant cDNAs were  
 selected and amplified by PCR reaction using vector region  
 3' primer. The PCR products were used as template for  
 synthesis of biotinylated single stranded RNA by in vitro  
 transcription reaction. The synthesized RNA probes were  
 hybridized with antisense single stranded cDNAs prepared  
 from original library and incubated with avidin-gel.  
 After removing DNA-RNA hybrids by centrifuge, the  
 subtracted cDNA libraries were constructed by  
 transformation of the remaining DNA into competent cells E.  
 coli TOP10F with electroporation method."

BASE COUNT 150 a 142 c 176 g 123 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,48e-105  
 Score: 938.00  
 Percent Similarity: 98.358  
 Best Local Similarity: 97.258  
 Query Match: 97.81%  
 DB: 14 Gaps: 0

US-10-082-502-19 (1-182) x BM832888 (1-591)

QY 1 MetLysGlyTyrPglYTrpLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
 |||||  
 DB 2 ATGAAGAAGGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 61  
 QY 21 ArgArgSerGlnAspLeuHscysGlyAlaCysArgAlaLeuValAspLeuLeuGlyTrp 40  
 |||||  
 DB 62 CGAGAGAGCGAGAGATCCACTGTGGAGCATGACAGGCGCTGTGTGATCACTAGAAATGG 121  
 QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 |||||  
 DB 122 GAAATTTGCCAGGTGACCCCAAGAACACATTCAGATGGATCTTCCGATCAATCA 181  
 QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
 |||||  
 DB 182 GATGGCAGCCACTGATGGTGGAGTGGCTTATGCCCTCAGAGGCCACCTCCACAGAG 241  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100  
 |||||  
 DB 242 CTGCTGGAGGAGATATGTGACCGATGAAAGGATGTGGGAAACAGATTGATCCTCCACC 301  
 QY 101 HisArgLysAsnTyrValArgValValSerArgAsnGlyGlySerGluLeuAspLeu 120  
 |||||  
 DB 302 CATGCCAAGAACTACGTACGTACTGTAGTGGCGCGAATGAGAAATCCAGTGAACCTGACCTA 361  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140

|||||  
Db 362 CAGGCAATCCGATGACTGACATATTACGGCACCCTCAATCTTGGCGTAGAGCAT 421  
Oy 141 VALGUGUUTYUGUASPGLUENLEUGLUPHESPESERATGUALAASPASNVALYS 160  
Db 422 GTGGAGGATAGAGAGTGAATCATTTGAAATCTTTCCCGAGAGCGTACAAATTTAA 481  
Oy 161 ASPYLSEUCYSERLYARGTHRASPLEUCYASPHISALALENHISARGSERHISAP 180  
Db 482 GACCAACTTTCAGTAGAGACAGATCTTTGTGACCATGCCCTGCACATATGCAATGAT 541  
Oy 181 GUAU 182  
Db 542 GAGCTA 547  
RESULT 12  
LOCUS B1712899 611 bp mRNA linear EST 11-MAR-2002  
DEFINITION 1698f09.y1 HR85 1slet Homo sapiens CDNA clone IMAGE:5085832 5'  
similar to TR:Q9Y2B0 Q9Y2B0 TYPE II MEMBRANE PROTEIN.; mRNA  
sequence.  
ACCESSION B1712899  
VERSION B1712899.1 GI:15688594  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 611)  
Mellon,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Keastner,K.,  
Lemisha,I., Searce,M., Brestelli,J., Grawohl,G., Clifton,S.,  
Hiller,L., Marra,M., Pape,D., Wylie,T., Martin,J., Bistola,A.,  
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas  
,M., Gibbons,M., McCann,R., Cole,R., Tsagarelisvilli,R., Williams,T.,  
Jackson,Y. and Bowers,Y.  
Endocrine Pancreas Consortium  
TITLE Unpublished (2000)  
JOURNAL Contact: Douglas Mellon, Klaus H. Kaestner, & Hiroshi Inoue  
COMMENT Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmellon@biohp.harvard.edu  
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center for information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@lm.wustl.edu)  
Seq primer: -40RP from GIBCO  
High quality sequence stop: 463.  
FEATURES  
source  
1. .611  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5085832"  
/clone\_lib="HR85 1slet"  
/tissue\_type="Purified pancreatic 1slet"  
/lab\_host="DH10B"  
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site:1;  
NotI; Site:2; XhoI; CDNA made by oligo-dt priming.  
Size-selected on agarose gel. Average insert size ~1kb. 5'  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Permutt Lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."

BASE COUNT 157 a 143 c 184 g 127 t  
ORIGIN  
Alignment Scores: 1.55e-105 Length: 611  
Pred. No.:

Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: 13 Gaps: 0  
US-10-082-502-19 (1-182) x B1712899 (1-611)  
Oy 1 MetlysglyttrpglytrpleuualaleuenuuglyvalleuenglythrAlatrrpAla 20  
Db 30 ATGAAGCGTGCGGTGGCTGGCCCTGCTTCTGGGGGCCCTGCGGAACGGCCGGGCT 89  
Oy 21 ArgArSerGlnAspLeuHiscysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
Db 90 CGAGAGCCAGCAGATCTCCACTGTGAGACATGACAGGCGCTGTGTGATGAATCAATG 149  
Oy 41 GUALEALARGVALASPROLYSLYSTRHLEIMETGLSERPHARGTLEASP 60  
Db 150 GAAATTTGCCAGGTGAGCCCAAGAAACCAATTCAGATGGATCTTCCGGATCAATCCA 209  
Oy 61 ASPGLYSERGLNSERVALVALGLUVALPROTYRALAARGSERGLUALAHISLEUTHRG 80  
Db 210 GATGGCAGCAGCTAGTGGTGGAGGTGGCTTATGCCCGTCAAGGCCACCTCAGAG 269  
Oy 81 LeuLeuGluGluValCysArgArgMetlysglyttrglygluGluIleAspProSerThr 100  
Db 270 CTGCTGGAGGAGATGTGTGACCGCATGAGAGATGTGGGAAACAGATTCATCTCCACC 329  
Oy 101 HISARGLYSANTTYVALARGVALASERARGANSGLYSERGLUENUSPLEU 120  
Db 330 CATCCAGAACACTACGTACGTGTACTGGCCGGAATGGAGATCCAGTGAATGGACCTA 389  
Oy 121 GlnGlyIleArgGlyLeaSPSerAspIleSerGlyThrLeuLyPheAlaCysGluSerIle 140  
Db 390 CAGGCAATCCGATGACTGACATATGTAGCGGCACCTCAAGTTGGCGTAGAGCAT 449  
Oy 141 VALGUGUUTYUGUASPGLUENLEUGLUPHESPESERATGUALAASPASNVALYS 160  
Db 450 GTGGAGGATAGAGAGTGAATCATTTGAAATCTTTCCCGAGAGCGTACAAATTTAA 509  
Oy 161 ASPYLSEUCYSERLYARGTHRASPLEUCYASPHISALALENHISARGSERHISAP 180  
Db 510 GACCAACTTTCAGTAGAGACAGATCTTTGTGACCATGCCCTGCACATATGCAATG 569  
Oy 181 GUAU 182  
Db 570 GAGCTA 575  
RESULT 13  
LOCUS BF344334 697 bp mRNA linear EST 22-NOV-2000  
DEFINITION 602014708P1 NCI\_CGAP\_Brn64 Homo sapiens CDNA clone IMAGE:4150512  
5', mRNA sequence.  
ACCESSION BF344334  
VERSION BF344334.1 GI:11291554  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 697)  
NIH-MGC http://mgc.ncl.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cga@ds-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov





QY 121 GINGLYTLEARGTLEASPERASPILESERGTYHTLEULYPHEALACYSGLUSERTIE 140  
DB 408 CAAGGATCCGATCGACATGATATACGGCACCCTCAAGTTCCGGTGAAGACAT 467  
QY 141 VALGLUGLUTYRGUASPLULEULIEGLUPHEPESERARGUALAASPAANYALLS 160  
DB 468 GTGGAGGATATGAGAGATGACATGATGATCTTTCCCGAGAGGCTACAAATGTTAA 527  
QY 161 ASPLYSLEUCYSESLYSARGTHRASPLEUCYASPHISALALEUHLISARGSERHISAP 180  
DB 528 GACAAACTTGCAGTAAAGCAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 587  
QY 181 GULU 182  
DB 588 GAGCTA 593  
RESULT 15  
LOCUS BG824549 742 bp mRNA linear EST 22-MAY-2001  
DEFINITION 602728454F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4868100 5',  
mRNA sequence.  
ACCESSION BG824549  
VERSION BG824549.1 GI:14172136  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 742)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-rt@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)  
DNA Sequencing by: NIH Intramural Sequencing Center  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ULNL at:  
http://image.llnl.gov  
Plate: L1CM1737 row: k column: 13  
High quality sequence stop: 740.  
Location/Qualifiers  
1. 742  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4868100"  
/clone\_1b="NIH\_MGC\_15"  
/tissue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: colon; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dt priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GCGACAGC(G). Size selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the Laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)"  
BASE COUNT 180 a 184 c 233 g 144 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,06e-105 Length: 742  
Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: 12 Gaps: 0  
US-10-082-502-19 (1-182) x BG824549 (1-742)  
QY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20

DB 141 ATGAAGAGCTGGGGTGGCTGGCCCTGCTTCGGGGGCCCTCGGAACCCCTGGCGCT 200  
QY 21 ARGARGSERGLASPLEUHLISCYSGLYALACYSARGALALEUVALASPLULEUGLUTP 40  
DB 201 CGGAGAGCCAGGATCTCCACTGTGGACATGACAGGCTCTGTGGATGAACTAGCAATGG 260  
QY 41 GUILLEALARGVALASPROLYSUTRILEGIMETGLYSERPHEARGLLEASPRO 60  
DB 261 GAAATTTGCCAGAGTGGAGCCCAAGAAAGACATTCAGATGGATCTTTCCGATCAATCA 320  
QY 61 ASPLYSERGLINSERVALALGLUVALPROTYRALAARGSERGLUALAHISLEUTHRGU 80  
DB 321 GATGGCAGCCAGTCAAGTGTGGAGGCGCTTATGCCCTCAGAGGCCACCTCACAGAG 380  
QY 81 LEULEUGLUGLVALCYSPARPMETLYSGIUTYRGYGLUGLNLLEASPROSETR 100  
DB 381 CTGCTGGAGAGATATGATGACCGGATGAAGAGATGAGGAACAGATTGATCTTCACCC 440  
QY 101 HISARGLYASNTYRVALARVALASERARGASNGLYGLUSERSERGLULEUASPLEU 120  
DB 441 CATCGCAAGACTAGCTAGCTAGTGGCCGGAATGGAATCCAGTGAATCGACTGGACCTA 500  
QY 121 GINGLYTLEARGTLEASPERASPILESERGTYHTLEULYPHEALACYSGLUSERTIE 140  
DB 501 CAAGGATCCGATGACATGACATGATATACGGCACCCTCAAGTTCCGGTGAAGACAT 560  
QY 141 VALGLUGLUTYRGUASPLULEULIEGLUPHEPESERARGUALAASPAANYALLS 160  
DB 561 GTGGAGGATATGAGAGATGACATGATGATCTTTCCCGAGAGGCTACAAATGTTAA 620  
QY 161 ASPLYSLEUCYSESLYSARGTHRASPLEUCYASPHISALALEUHLISARGSERHISAP 180  
DB 621 GACAAACTTGCAGTAAAGCAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 680  
QY 181 GULU 182  
DB 681 GAGCTA 686

Search completed: March 31, 2003, 04:51:08  
Job time : 1169.85 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: March 31, 2003, 07:02:46 ; Search time 55.0233 Seconds

(without alignments)  
2614.757 Million cell updates/sec

Title: US-10-082-502-19

Perfect score: 959

Sequence: 1 MKGWMLALLGLVLTGTA.....LCRKRTDLCDAHRSNDEL 182

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 574371 segs, 425486471 residues  
Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame\_plus.model -DEV=x1p  
-Q=/cgn2\_1/USPRO\_pool/US0082502/runat\_24032003\_135103\_6870/app\_query.fasta\_1.1308  
-DB=Published\_Applications\_NA -QFMT=fastcap -SUFFIX=trmp -MINMATCH=0.1  
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR=SCORE-pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10082502.@CGN.1.1.145 @runat\_24032003\_135103\_6870  
-NCP=6 -ICPU=3 -NO\_XLPEX -NO\_MMAP -LARGEODERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published\_Applications\_NA:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	938	97.8	814 10 US-09-864-864-309	Sequence 309, App
2	938	97.8	1210 9 US-10-028-072-173	Sequence 173, App
3	938	97.8	1210 9 US-10-121-049-173	Sequence 173, App
4	938	97.8	1210 9 US-10-123-904-173	Sequence 173, App

5	938	97.8	1210 9	US-10-140-470-173	Sequence 173, App
6	938	97.8	1210 9	US-10-175-746-173	Sequence 173, App
7	938	97.8	1210 9	US-10-176-918-173	Sequence 173, App
8	938	97.8	1210 9	US-10-176-921-173	Sequence 173, App
9	938	97.8	1210 9	US-10-137-865-173	Sequence 173, App
10	938	97.8	1210 9	US-10-140-474-173	Sequence 173, App
11	938	97.8	1210 9	US-10-142-431-173	Sequence 173, App
12	938	97.8	1210 9	US-10-143-114-173	Sequence 173, App
13	938	97.8	1210 9	US-10-140-002-173	Sequence 173, App
14	938	97.8	1210 9	US-10-142-419-173	Sequence 173, App
15	938	97.8	1210 9	US-10-123-262-173	Sequence 173, App
16	938	97.8	1210 9	US-10-142-423-173	Sequence 173, App
17	938	97.8	1210 9	US-10-121-050-173	Sequence 173, App
18	938	97.8	1210 9	US-10-141-755-173	Sequence 173, App
19	896	93.4	657 9	US-09-822-846-398	Sequence 398, App
20	262	27.3	564 10	US-09-864-864-228	Sequence 228, App
21	136.5	14.2	1457 10	US-09-846-5738-10	Sequence 10, App
22	136.5	14.2	1470 9	US-10-036-041-58	Sequence 58, App
23	136.5	14.2	1470 9	US-10-035-853-58	Sequence 58, App
24	136.5	14.2	1470 9	US-10-174-590-479	Sequence 479, App
25	136.5	14.2	1470 9	US-10-176-758-479	Sequence 479, App
26	136.5	14.2	1470 9	US-10-175-737-479	Sequence 479, App
27	136.5	14.2	1470 9	US-10-173-706-479	Sequence 479, App
28	136.5	14.2	1470 9	US-10-175-728-479	Sequence 479, App
29	136.5	14.2	1470 9	US-10-175-732-479	Sequence 479, App
30	136.5	14.2	1470 9	US-10-176-482-479	Sequence 479, App
31	136.5	14.2	1470 9	US-10-176-757-479	Sequence 479, App
32	136.5	14.2	1470 9	US-10-176-913-479	Sequence 479, App
33	136.5	14.2	1470 9	US-10-180-553-479	Sequence 479, App
34	136.5	14.2	1470 9	US-10-180-557-479	Sequence 479, App
35	136.5	14.2	1470 9	US-09-931-836-58	Sequence 58, App
36	136.5	14.2	1470 9	US-10-173-700-479	Sequence 479, App
37	136.5	14.2	1470 9	US-10-174-572-479	Sequence 479, App
38	136.5	14.2	1470 9	US-10-174-582-479	Sequence 479, App
39	136.5	14.2	1470 9	US-10-174-588-479	Sequence 479, App
40	136.5	14.2	1470 9	US-10-175-733-479	Sequence 479, App
41	136.5	14.2	1470 9	US-10-175-739-479	Sequence 479, App
42	136.5	14.2	1470 9	US-10-175-743-479	Sequence 479, App
43	136.5	14.2	1470 9	US-10-175-743-479	Sequence 479, App
44	136.5	14.2	1470 9	US-10-176-488-479	Sequence 479, App
45	136.5	14.2	1470 9	US-10-176-492-479	Sequence 479, App

## ALIGNMENTS

RESULT 1  
US-09-864-864-309  
Sequence 309, Application US/09864864  
Patent No. US20020102679A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Mitchell, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Secrist, Heather  
APPLICANT: Lodes, Michael J.  
APPLICANT: Algate, Paul A.  
APPLICANT: Filing, Steve P.  
APPLICANT: Benson, Jane  
APPLICANT: Benson, Darin R.  
APPLICANT: Carter, Darick  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.523  
CURRENT APPLICATION NUMBER: US/09/864, 864  
CURRENT FILING DATE: 2001-05-23  
NUMBER OF SEQ ID NOS: 341  
SOFTWARE: Corixa Invention Disclosure Database  
SEQ ID NO 309  
LENGTH: 814  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-864-864-309

Alignment Scores:

Pred. No.: 5,84e-111 Length: 814  
Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: 10 Gaps: 0

US-10-082-502-19 (1-182) x US-09-864-864-309 (1-814)

```

Oy 1 Metlysglytrpglytrplealaaleuenglyvalleuenglythralatpala 20
    |||||||
Db 145 ATGAAAGCTGGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 204
Oy 21 Argatserglnasrpleuhalasysglvalacysatgalaaleuvalaspluueu 40
    |||||||
Db 205 CGAGAGACCCAGATCTCCAGCTGGAGACATGAGAGCTGGTGGATGAGATCAATGG 264
Oy 41 Glutlealaatgvalasrprolyalsthrileglnmetglsrphatrgllasrpro 60
    |||||||
Db 265 GAAATGGCCAGGTGGAGCCCAAGAGACATTCAGATGGATGATCTTCCGATCAATCA 324
Oy 61 Aspglyserglnasrvalalvala1protyralaargserglualahisrthr 80
    |||||||
Db 325 GATGGACACCCAGTGGTGGAGGTGGCTTATGGCTTATGGCTTATGGCTTATGGCT 384
Oy 81 Leuleuglgluvalcysasrpargmetlysglytrgluglnleasrproserthr 100
    |||||||
Db 385 CTGCTGGAGAGATATGTGACCGGATGAGAGATGAGAGATGAGATGATGATGATGAT 444
Oy 101 Hisarglysaanttyrvalargvalalaserarganglyglusersegluleasrpleu 120
    |||||||
Db 445 CATCCCAAGAACTAGTGTGTCTAGTGGCCGGAATGAGATCCAGTCACTGACCTGACCTA 504
Oy 121 Glnglylalarglyleasrserasrplesergylthreulysphela1cysgluserile 140
    |||||||
Db 505 CAAGGCATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 564
Oy 141 Valglutlytrgluasrpluleuilegluphepserarglualaasrparvallys 160
    |||||||
Db 565 GTGGAGAAATGAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 624
Oy 161 Asrlyasleucysserlyarglthrasrpleucysasrphalaleuhalasrserthasr 180
    |||||||
Db 625 GACAACTTTCAGTAAAGCAAGATCTTGTGACCATGCTGACATGATGATGATGATGAT 684
Oy 181 Gluleu 182
    |||||||
Db 685 GAGCTA 690

```

RESULT 2

US-10-028-072-173  
Sequence 173, Application US/10028072  
Publication No. US2003000431A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Flivaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey J.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang

TITLE OF INVENTION:

FILE REFERENCE:

CURRENT APPLICATION NUMBER: US/10/028, 072  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: 60/049911  
PRIOR FILING DATE: 1997-06-18  
PRIOR APPLICATION NUMBER: 60/056974  
PRIOR FILING DATE: 1997-08-26  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059115  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059117  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059122  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059184  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059352  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059588  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059836  
PRIOR FILING DATE: 1997-09-24  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062285  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062287  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062814  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/062816  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063045  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063082  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/063127  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063327  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063329  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063550  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063561  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063704  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063733  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063735  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063738  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063755  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064248  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/064809  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065846  
PRIOR FILING DATE: 1997-11-17  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/066453  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066511

PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/066770  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/069212  
 PRIOR FILING DATE: 1997-12-11  
 PRIOR APPLICATION NUMBER: 60/069278  
 PRIOR FILING DATE: 1997-12-11  
 PRIOR APPLICATION NUMBER: 60/069334  
 PRIOR FILING DATE: 1997-12-11  
 PRIOR APPLICATION NUMBER: 60/069694  
 PRIOR FILING DATE: 1997-12-16  
 PRIOR APPLICATION NUMBER: 60/072320  
 PRIOR FILING DATE: 1998-01-23  
 PRIOR APPLICATION NUMBER: 60/073612  
 PRIOR FILING DATE: 1998-02-04  
 PRIOR APPLICATION NUMBER: 60/074086  
 PRIOR FILING DATE: 1998-02-09  
 PRIOR APPLICATION NUMBER: 60/074092  
 PRIOR FILING DATE: 1998-02-09  
 PRIOR APPLICATION NUMBER: 60/077791  
 PRIOR FILING DATE: 1998-03-12  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/079294  
 PRIOR FILING DATE: 1998-03-25  
 PRIOR APPLICATION NUMBER: 60/079663  
 PRIOR FILING DATE: 1998-02-27  
 PRIOR APPLICATION NUMBER: 60/079728  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/080165  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081695  
 PRIOR FILING DATE: 1998-04-14  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081818  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082999  
 PRIOR FILING DATE: 1998-04-24  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084627  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084637  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/085149  
 PRIOR FILING DATE: 1998-05-12  
 PRIOR APPLICATION NUMBER: 60/085323  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085338  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085339  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085579  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085697  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085704  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/086414  
 PRIOR FILING DATE: 1998-05-22  
 PRIOR APPLICATION NUMBER: 60/086430  
 PRIOR FILING DATE: 1998-05-22  
 PRIOR APPLICATION NUMBER: 60/087106  
 PRIOR FILING DATE: 1998-05-28

PRIOR APPLICATION NUMBER: 60/088026  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088730  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088741  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088810  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088858  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090445  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090538  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090863  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/091360  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091519  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091982  
 PRIOR FILING DATE: 1998-07-07

Alignment Scores:  
 Pred. No.: 1,01e-110 Length: 1210  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3  
 Query Match: 97.81% Indels: 0  
 DB: 9 Gaps: 0

US-10-082-502-19 (1-182) x US-10-028-072-173 (1-1210)

QY 1 MetLysGLYTrpGLYTrpLeuLLeuLLeuGLYValLeuLLeuGLYThrAlaTrpAla 20  
 DB 543 ATGAAGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 602  
 QY 21 ArgArgSerGlnAspLeuHIsCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 DB 603 CGAGAGAGCCAGGATCTCCAGCTGAGCATGAGGAGGCTGTGTGATACATAGATG 662  
 QY 41 GluLLeuAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgLLeuAspPro 60  
 DB 663 GAAATTTGCCAGGTGAGCCCAAGAACCAATTCAGATGATGATCTTCGGATCAATCCA 722  
 QY 61 AspGlySerGlnSerValValGluValProTyraLeuArgSerGluAlaHisLeuTrpGlu 80  
 DB 723 GATGCGAGCATGACTGCTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 782  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 100  
 DB 783 CTGCTGAGAGAGATATGATCCGAGATGAGAGATGAGAGATGAGAGATGAGAGATG 842  
 QY 101 HisArgLysAspTyraValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 120  
 DB 843 CATCGAAGAACTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 902  
 QY 121 GlnGlyLLeuArgLLeuAspSerAspLLeuSerGlyThrLeuValPheAlaCysGluSerLle 140  
 DB 903 CAAGGCATCCGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 962

QY	1	MetuysGLYTPRGLYTPRPLEuAlaLeuLeuLeuAlaLeuLeuGLYTHPAlaTPAla	20
Db	543	ATGAAGAGCGTGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	602
QY	21	ArgArgSerGlnAspLeuHisCysGlyValAcysArgAlaLeuValaAspGluLeuGluTPR	40
Db	603	CGGAGGAGCCAGGATCTCCACTGTGGAGCATGCAGGGCTGTGGTGATGAACCTGAAATGG	662
QY	41	GluIleuLeuAlaArgValaAspProGluLeuThrIleGluMetGlySerPheAlaGluIleAspPro	60
Db	663	GAAATTCGCCAGGGGAGACCCCAAGAMACACATTCAGATGGATCTTTCCGAGATCAATCCA	722
QY	61	AspGlySerGlnSerValaGluValaProCysAlaIaArgSerGluAlaHisLeuThrGlu	80
Db	723	GATGGACGCCAGTCAAGTGGTGGAGAGCTGCTTAAATCCCGGCTCAAGAGGCCACCTACAGAG	782

qy 1 MetLysGlyTyrIpsIyTyrPheuAlaLeuLeuLeuGlyValLeuLeuGlyThralaIrrPaLa 20  
 |||||  
 Db 543 ATGAAGGCTGGGGTTGGCTGGCCCTGCTTCTGGGGGCCCTGCTGGGAACCGCTGGGCT 6022

QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
Db 603 CGAGAGCCAGGATCTCTCAGCTGGAGCATGACAGGCTCTGGATGAACCTAGCAATGG 662  
QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
Db 663 GAAATGGCCAGGTGGACCCCAAGAACCTTCAGATGGATCTTTCCGATCAATCCA 722  
QY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
Db 723 GATGGACACCCAGTCAGTGTGGAGTGGCTTATGCCCCCTCAGAGCCCACTCAGACAG 782  
QY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
Db 783 CTGCTGGAGGAGATATGTGACCGAGTGAAGAGATGGGAAACAGATTGATCTTCCACC 842  
QY 101 HisArgLysAsnTyrValaArgValaValSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
Db 843 CATCGCAAGAACTAGTACGTAGTGGCCCGGAATGGAAATCCAGTGAACCTGACCTA 902  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 903 CAAGGCATCCGAAATCGACTCAGATATTAGCGCACCTCAAGTTTGGGTGGAGACATTT 962  
QY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
Db 963 GTGAGGAGAAATCGAGATGAACCTATTGATCTTTCCCGAGAGGCTGACAAATGTTAAA 1022  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
Db 1023 GACAAACTTTGACAGTAAAGCAAGATCTTGTGACCATGGCCCTGCACATATGCAATGAT 1082  
QY 181 GluLeu 182  
Db 1083 GAGCTA 1088

RESULT 5  
US-10-140-470-173  
; Sequence 173, Application US/10140470  
; Publication No. US20030022331A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C160  
; CURRENT APPLICATION NUMBER: US/10/140,470  
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-140-470-173

Alignment Scores: 1.01e-110 Length: 1210  
Pred. No.: 1210

Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-19 (1-182) x US-10-140-470-173 (1-1210)

QY 1 MetLysGlyTyrPglTyrPLeuAlaLeuLeuGluValLeuGluGlyThrAlaTrpAla 20  
Db 543 ATGAAGCTGGGGTGGCTGGCCCTCTCTGCGGAGCATGACGGGCTGTGGATGAACATAAATGG 602  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
Db 603 CGAGAGCCAGGATCTCTCAGCTGGAGCATGACAGGCTCTGGATGAACATAAATGG 662  
QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
Db 663 GAAATGGCCAGGTGGACCCCAAGAACCTTCAGATGGATCTTTCCGATCAATCCA 722  
QY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
Db 723 GATGGACACCCAGTCAGTGTGGAGTGGCTTATGCCCCCTCAGAGCCCACTCAGACAG 782  
QY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
Db 783 CTGCTGGAGGAGATATGTGACCGAGTGAAGAGATGGGAAACAGATTGATCTTCCACC 842  
QY 101 HisArgLysAsnTyrValaArgValaValSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
Db 843 CATCGCAAGAACTAGTACGTAGTGGCCCGGAATGGAAATCCAGTGAACCTGACCTA 902  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 903 CAAGGCATCCGAAATCGACTCAGATATTAGCGCACCTCAAGTTTGGGTGGAGACATTT 962  
QY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
Db 963 GTGAGGAGAAATCGAGATGAACCTATTGATCTTTCCCGAGAGGCTGACAAATGTTAAA 1022  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
Db 1023 GACAAACTTTGACAGTAAAGCAAGATCTTGTGACCATGGCCCTGCACATATGCAATGAT 1082  
QY 181 GluLeu 182  
Db 1083 GAGCTA 1088

RESULT 6  
US-10-175-746-173  
; Sequence 173, Application US/10175746  
; Publication No. US20030027270A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C353

```

: CURRENT APPLICATION NUMBER: US/10/175,746
: CURRENT FILING DATE: 2002-06-19
: Prior Application removed - See file wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 173
: LENGTH: 1210
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-175-746-173

Alignment Scores:
Pred. No.: 1,01e-110      Length: 1210
Score: 938.00             Matches: 177
Percent Similarity: 98.35% Conservative: 2
Best Local Similarity: 97.25% Mismatches: 3
Query Match: 97.81%       Indels: 0
DB: 9                      Gaps: 0

US-10-082-502-19 (1-182) x US-10-175-746-173 (1-1210)

QY 1 MetLysGIYTRPGIYTRPLeuAlaLeuLeuGIYValLeuGIYThrAlaTrpAla 20
D 543 ATGAAGCGCTGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 602
QY 21 ArgArgSerGIaAPLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40
D 603 CGGAGGACCCGAGATCTCCAGTGGACATGCAGAGGCTCTGGTGATGAACCTGAATGG 662
QY 41 GluIleAlaArgValaAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60
D 663 GAATTGCCCGAGGTGGACCCCAAGAACATTCAGATGGGATCTTCCGATCAATCCA 722
QY 61 AspGlySerGlnSerValaIleGluValaProTyraAlaArgSerGluAlaHISLeuThrglu 80
D 723 GATGGCACCCGAGTGGTGGAGTGGCTTATGCCCTCAGAGGCCACCTCAGAGAG 782
QY 81 LeuLeuGluGluValaCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 100
D 783 CTGCTGGAGGAGATATGTGACCCGATGAGAGATGGGAAACAGATGATCTTCCACC 842
QY 101 HisArgLysAsnTyraValaArgValaIleSerArgAsnGlyLysSerGluLeuAspLeu 120
D 843 CATGCAAGAACTACGTACGTAGTGGGCGGAGATGAGAACTCAGTAACCTGAGCTA 902
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
D 903 CAAGCAATCCCAATCGATCAGATATTAGCGCACCCCAAGTTTGCCTGTGAGAGCT 962
QY 141 ValGluGluTyrglyLysAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160
D 963 GTGAGAGCATCGAGAGATGAACATTCATTTCTTTCCGAGAGGCTGACAAATGTTAA 1022
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHISAlaLeuHISArgSerHISAsp 180
D 1023 GACAAACTTTCAGATGAAGCGAAGAGATCTTGTGACCATGCTGCAATATCGATGAT 1082
QY 181 GluLeu 182
D 1083 GAGCTA 1088

RESULT 7
US-10-176-918-173
: Sequence 173, Application US/10176918
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerltsen, Mary E.
: APPLICANT: Goddard, Audrey
```

```

: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tamas, Daniel
: APPLICANT: Matanade, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P33081C382
: CURRENT APPLICATION NUMBER: US/10/176,918
: CURRENT FILING DATE: 2002-06-20
: Prior Application removed - See file wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 173
: LENGTH: 1210
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-176-918-173

Alignment Scores:
Pred. No.: 1,01e-110      Length: 1210
Score: 938.00             Matches: 177
Percent Similarity: 98.35% Conservative: 2
Best Local Similarity: 97.25% Mismatches: 3
Query Match: 97.81%       Indels: 0
DB: 9                      Gaps: 0

US-10-082-502-19 (1-182) x US-10-176-918-173 (1-1210)

QY 1 MetLysGIYTRPGIYTRPLeuAlaLeuLeuGIYValLeuGIYThrAlaTrpAla 20
D 543 ATGAAGCGCTGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 602
QY 21 ArgArgSerGIaAPLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40
D 603 CGGAGGACCCGAGATCTCCAGTGGACATGCAGAGGCTCTGGTGATGAACCTGAATGG 662
QY 41 GluIleAlaArgValaAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60
D 663 GAATTGCCCGAGGTGGACCCCAAGAACATTCAGATGGGATCTTCCGATCAATCCA 722
QY 61 AspGlySerGlnSerValaIleGluValaProTyraAlaArgSerGluAlaHISLeuThrglu 80
D 723 GATGGCACCCGAGTGGTGGAGTGGCTTATGCCCTCAGAGGCCACCTCAGAGAG 782
QY 81 LeuLeuGluGluValaCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 100
D 783 CTGCTGGAGGAGATATGTGACCCGATGAGAGATGGGAAACAGATGATCTTCCACC 842
QY 101 HisArgLysAsnTyraValaArgValaIleSerArgAsnGlyLysSerGluLeuAspLeu 120
D 843 CATGCAAGAACTACGTACGTAGTGGGCGGAGATGAGAACTCAGTAACCTGAGCTA 902
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
D 903 CAAGCAATCCCAATCGATCAGATATTAGCGCACCCCAAGTTTGCCTGTGAGAGCT 962
QY 141 ValGluGluTyrglyLysAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160
D 963 GTGAGAGCATCGAGAGATGAACATTCATTTCTTTCCGAGAGGCTGACAAATGTTAA 1022
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHISAlaLeuHISArgSerHISAsp 180
D 1023 GACAAACTTTCAGATGAAGCGAAGAGATCTTGTGACCATGCTGCAATATCGATGAT 1082
QY 181 GluLeu 182
D 1083 GAGCTA 1088

RESULT 8
```



```

US-10-176-921-173
; Sequence 173, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvarolf, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C28
; CURRENT APPLICATION NUMBER: US/10/176,921
; PRIOR FILING DATE: 2002-06-20
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-921-173

Alignment Scores:
Pred. No.: 1 01e-110
Score: 938.00
Percent Similarity: 98.35%
Best Local Similarity: 97.25%
Query Match: 97.81%
Matches: 1210
Conservative: 2
Mismatch: 3
Indels: 0
Gaps: 0

US-10-082-502-19 (1-182) x US-10-176-921-173 (1-1210)
QY 1 Metlysglytrpdlrpleuallaleuileuileuglythraltatpala 20
Db 543 ATGAAAGGCTGGGTTGGCTGGCCCTCTCTGGGGGCGCTGCGGAAACCGCTGGGCT 602
QY 21 ArgArgserGlnaspLeuHIScysGlyAlaCysArgAlaLeuValaspGluLeuGluTrp 40
Db 603 CGGAGGAGCCAGATCTCCACTGTGGAGCATGCGAGGCTGTGGTGAAGTGAAGATGG 662
QY 41 GluIleAlaArgValaspProLysThrIleGlnMetGlySerPheArgIleasnPro 60
Db 663 GAAATTCGCCAGGTGGAGCCCAAGAGACCATTCAGATGGGATCTTCCCGATCAATCCA 722
QY 61 AspGlyserGlnserValValGluValProTyrAlaArgserGluAlaHISLeuThrGlu 80
Db 723 GATGGCAGCCAGTCAGTGTGGAGTGCCTTATGCCCGCTCAGAGGCCACCTCAGAG 782
QY 81 LeuLeuGluIleValCysAspArgMetLysGluTyrGlyGluGlnIleaspProSerThr 100
Db 783 CTCTGTGAGAGATATGTACCGGATGAGAGATGTGGGAAACGATGTGATCTTCCACC 842
QY 101 HisArgLysAsnTyrValAlaArgValValSerArgAsnGlyLysSerGluLeuAspLeu 120
Db 843 CATCGCAAGACATACGATGATGTGGGCGGAGATGAGATCAATGAACTGACCTA 902
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 903 CAAGGCAATCCGAATCGACTCAGATATTAAGCGCACCTCAAGTTGGGTGAGAGCAT 962
QY 141 ValGluIleTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160

```

```

Db 963 GTGGAGGATACGAGATGACATTCATTTTCCCGAGGCTGCAATGTATAA 1022
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHISAlaLeuHISArgSerHISAsp 180
Db 1023 GACAACTTTCGAGTAAAGCAAGATCTTTGTGACATGCGCTCGCAGATATGCAATGAT 1082
QY 181 GluLeu 182
Db 1083 GACCTA 1088

RESULT 9
US-10-137-865-173
; Sequence 173, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvarolf, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C154
; CURRENT APPLICATION NUMBER: US/10/137,865
; PRIOR FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-137-865-173

Alignment Scores:
Pred. No.: 1 01e-110
Score: 938.00
Percent Similarity: 98.35%
Best Local Similarity: 97.25%
Query Match: 97.81%
Matches: 1210
Conservative: 2
Mismatch: 3
Indels: 0
Gaps: 0

US-10-082-502-19 (1-182) x US-10-137-865-173 (1-1210)
QY 1 Metlysglytrpdlrpleuallaleuileuileuglythraltatpala 20
Db 543 ATGAAAGGCTGGGTTGGCTGGCCCTCTCTGGGGGCGCTGCGGAAACCGCTGGGCT 602
QY 21 ArgArgserGlnaspLeuHIScysGlyAlaCysArgAlaLeuValaspGluLeuGluTrp 40
Db 603 CGGAGGAGCCAGATCTCCACTGTGGAGCATGCGAGGCTGTGGTGAAGTGAAGATGG 662
QY 41 GluIleAlaArgValaspProLysThrIleGlnMetGlySerPheArgIleasnPro 60
Db 663 GAAATTCGCCAGGTGGAGCCCAAGAGACCATTCAGATGGGATCTTCCCGATCAATCCA 722
QY 61 AspGlyserGlnserValValGluValProTyrAlaArgserGluAlaHISLeuThrGlu 80
Db 723 GATGGCAGCCAGTCAGTGTGGAGTGCCTTATGCCCGCTCAGAGGCCACCTCAGAG 782
QY 81 LeuLeuGluIleValCysAspArgMetLysGluTyrGlyGluGlnIleaspProSerThr 100

```

```
|||||
Db 793 CTGCTGGAGGAGATATGACCGCATGAGAGAGATGGGGAACACATGATCTCCACAC 842
OY 101 HsArgLysAsnTyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 120
Db 843 CATCCCAAGAACTACGCTGCTGCTAGTGGCGCGGAATGGAATCCAGTGAACCTGACCTA 902
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 903 CAAGCATCCGAAATCGATCAGATATGTAGCGGACCTCAAGTTGGCTGTGAGAGCAATT 962
OY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160
Db 963 GTGGAGGAATACGAGATGAACTCATTAATCTTTCCCGAGAGGCGACAAATGTTAAA 1022
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180
Db 1023 GACAAACTTTGCAGTAACGAAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 1082
OY 181 GluLeu 182
Db 1083 GAGCTA 1088

RESULT 10
US-10-140-474-173
; Sequence 173, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-474-173

Alignment Scores:
Pred. No.: 1,01e-110 Length: 1210
Score: 938.00 Matches: 177
Percent Similarity: 98.35% Conservative: 3
Best Local Similarity: 97.25% Mismatches: 2
Query Match: 97.81% Indels: 0
Gaps: 0
DB: 9

US-10-082-502-19 (1-182) x US-10-140-474-173 (1-1210)
OY 1 MetLysGlyTyrGlyTyrPheAlaLeuLeuLeuGlyValLeuLeuGlyThrAlaTyrPala 20
Db 543 ATGAAGAGCTGGGGTGGCTGGCTGCTGGGGGGCCCTGCTGGGAACCGCGCTGGGCT 602
OY 21 ArgArgSerGlnAspLeuHisCysGlyValAcysArgAlaLeuValAspGluLeuGluTyrP 40
```

```
|||||
Db 603 CGGAGAGACCCAGGATCTCCACCTGTGAGACATGACAGGCTGTGTGGATGAACTAGAAATGG 662
OY 41 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60
Db 663 GAAATTTGCCAGGTGGAGCCCAAGAAAGACATTCAGATGGAGGATCTTCCGGATCAATCA 722
OY 61 AspGlySerGlnSerValValGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 80
Db 723 GATGCGACCGCATGCTAGTGTGGAGGTGCTTATGCCCTGTAGAGGCCACCTGCACAGAG 782
OY 81 LeuLeuGluGluValCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100
Db 783 CTGCTGGAGGAGATATGACCGCATGAGAGATGTGGGGAACACATGATCTCCACAC 842
OY 101 HsArgLysAsnTyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 120
Db 843 CATCCCAAGAACTACGCTGCTGCTAGTGGCGCGGAATGGAATCCAGTGAACCTGACCTA 902
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 903 CAAGCATCCGAAATCGATCAGATATGTAGCGGACCTCAAGTTGGCTGTGAGAGCAATT 962
OY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160
Db 963 GTGGAGGAATACGAGATGAACTCATTAATCTTTCCCGAGAGGCGACAAATGTTAAA 1022
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180
Db 1023 GACAAACTTTGCAGTAACGAAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 1082
OY 181 GluLeu 182
Db 1083 GAGCTA 1088

RESULT 11
US-10-142-431-173
; Sequence 173, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-431-173

Alignment Scores:
Pred. No.: 1,01e-110 Length: 1210
Score: 938.00 Matches: 177
```

```
Percent Similarity: 98.35%
Best Local Similarity: 97.25%
Query Match: 97.81%
DB: 9
```

```
Conservative: 2
Mismatches: 3
Indels: 0
Gaps: 0
```

US-10-082-502-19 (1-182) x US-10-142-431-173 (1-1210)

QY	1	MelysgllytRPGlytRPLeuAlaIleuIleuGluValIleuIleuGllythrAlaTrpAla	20
Dd	543	ATGAAGCGCTGGGGTTCGTGGCCCTCCTCTTGCGGGCCCTGTGGAAACCCCCTGGGGT	602
QY	21	ArgHrSerglnAspIleuHisCysgylAlaCysArgAlaLeuValAspIleuGluTrp	40
Dd	603	CAGAGAGACCAGATCTCCACTGTGGAGCATGGAGGGCTGTGGTAATCAATTAAATGG	662
QY	41	GluIleuLalaArgValAspProlysLystrIleGlmMetGlySerPheArgIleasnPro	60
Dd	663	GAATTTCGCCAGGTGGAGCCCCAAGAAGACCATTCAATGAGATCTTTCCGGATCAATCCA	722
QY	61	AspIlySergInserValValGluValProTyraIlaArgSergluuAlaHisuehrGlu	80
Dd	723	GATGGCACCCAGTCACTAGTGTGGAGGTGCCTTAGGCCCTCCAGAGGCCCACTCCACAAG	782
QY	81	LeuDeuGluGluValCysAspArgMetLysGluTyrglGluAlileaspProSerThr	100
Dd	783	CTGCTGGAGGAGATATGTGTGACGGAGTGAAGAGATAGGGGAACAGATTGATCTTCCACC	842
QY	101	HisArgLysAsnTyraValaArgValValSerArgasnglyGluSerSergIleuAspIeu	120
Dd	843	CATGGCAAGAACATCACTACTGTAGTGGGCCCGGAAATGAGAAATCCAGTAACTGAGACCTA	902
QY	121	GlnglyIleArgIleaspSerAspIleSergIlythrLeuLysPheAlaCysGluSerIle	140
Dd	903	CMAAGCATCCGAATCGACTCAGATATTTACGGCACCCCTAAATTTTTCCGGTAGAGACATT	962
QY	141	ValGluGluTyrglunAspGluIleuIleGluPhePheSerArgGluAlaAspAsnValLys	160
Dd	963	GTTGAGAGAAATACGAGGATGAACATTTGAATCTTTTCCCGAGAGGCTGACAAATTATAA	1022
QY	161	AspLysIleucySerLysArgthrAspLeuCysAspHisAlaLeuHisArgSerHisAsp	180
Dd	1023	GACAAACTTTTGACAGTAAAGCAGACAGATCTTTGTGACACATGACCCCTGCACATATGCGATGAT	1082
QY	181	GluIeu 182	
Dd	1083	GAGCTA 1088	

```

RESULT 12
US-10-143-114-173
: Sequence 173, Application US/10143114
: Publication No. US2003036180A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P33303RIC211
: CURRENT APPLICATION NUMBER: US/10/143,114

```

```

; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173

```

### Alignment Scores

Pred. No.:	1,01e-110	Length:	1210
Score:	938.00	Matches:	177
Percent Similarity:	98.358	Conservative:	2
Best Local Similarity:	97.258	Mismatches:	3
Query Match:	97.818	Indels:	0
DB:	9	Gaps:	0

US-10-082-502-19 (1-182) x US-10-143-114-173 (1-1210)

QY	1	MetLysIlyrPglYrPLeuAlaLeuIleuLeuGluValIleuLeuGlyThrAlaPrla	20
Db	543	ATGAAAGCTGGGGTTGGCTGGCCCTGCTTCTGGGGCCCTGCTGGAAACCCCTGGGCT	602
QY	21	ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspIleuGluTrp	40
Db	603	CGAGAGACCGAGGATCTCCACTGTGGAGCATGCAAGGGCTGTGGTGAATCACTAAATGG	662
QY	41	GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro	60
Db	663	GAATTCTCCAGGTGGACCCCAAGAGACCATTCAAGTGGATCTTCCGGATCAATCCA	722
QY	61	AspIlySerGlnSerValValGluValProTyAlaArgSerGluAlaHisLeuThrGlu	80
Db	723	GATGGACCGACGATCAGTGGTGGAGGTGCTTATGCCGCTCAAGAGCCCACTTCACAG	782
QY	81	LeuLeuGluGluValCysAspArgMetLysGluTyGlyGluGlnIleAspSerThr	100
Db	783	CTGCTGAGAGAGATATGTGACCGGATGAAGAGATAGGGGAACAGTTATCTTCCACAC	842
QY	101	HisArgLysAsnIlyrValAlaArgValValSerArgAsnGlyLysSerSerGluAsnLeu	120
Db	843	CATGGCAAGACACGATCACTGTAGTGGGCGGAAATGAGAAATCCAGTCAATGGACCTA	902
QY	121	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle	140
Db	903	CAGAGCATCCGAATCGACTAGATATTAGGGCACCTCAAACTTTGGCGGTAGAGCAATT	962
QY	141	ValGluGluTrpGluLysAspGluLeuIleGluPhePheSerArgGluAlaAspArgValLys	160
Db	963	GTTGAGGAATACAGCGATGAACTATTGAATCTTTTCCGAGAGGCTGACAAATTGAAA	1022
QY	161	AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp	180
Db	1023	GACAAACTTTGCAGTAAAGCAACAGATCTTTGTGACATGACCTGCACATATGCGATGAT	1082
QY	181	GluLeu 182	
Db	1083	GAGCTA 1088	

RESULT 13  
 US-10-140-002-173  
 Sequence 173, Application US/10140002  
 Publication NO. US20030037623A1  
 GENERAL INFORMATION:  
 APPLICANT: Baker, Kevin P.  
 APPLICANT: Beresini, Maureen  
 APPLICANT: DeForge, Laura  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Filvaroff, Ellen  
 APPLICANT: Go, Wei-Qiang  
 APPLICANT: Gertsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.

```

? APPLICANT: Gurney,Austin L.
? APPLICANT: Sherwood,Steven
? APPLICANT: Smith,Victoria
? APPLICANT: Stewart,Timothy A.
? APPLICANT: Tumas,Daniel
? APPLICANT: Matanabe,Colin K
? APPLICANT: Wood,William
? APPLICANT: Zhang, Zemin
? TITLE OF INVENTION: ACIDS ENCODING THE SAME
? FILE REFERENCE: P33030R1C59
? CURRENT APPLICATION NUMBER: US/10/140,002
? PRIORITY FILING DATE: 2002-05-06
? Prior Application removed - See Palm or File Wrapper
? NUMBER OF SEQ ID NOS: 550
? SEQ ID NO 173
? LENGTH: 1210
? TYPE: DNA
? ORGANISM: Homo Sapien
? OS-10-140-002-173

```

Db 963 GTGAGAGATACAGATCACTCATTCATTTCCCGAGAGCTGACATATCTTAA 1022  
 QY 161 AsplysleuCySerlyArgThrAspleuCyAspHisAlaLeuHisArgSerHisAsp 180  
 Db 1023 GACAACTTTGCAGTACGAGACAGATCTTTGTGACCATGCCCTGCACATATCGATGAT 1082  
 QY 181 Glueu 182  
 Db 1083 GAGCTA 1088

## RESULT 15

US-10-123-262-173  
 : Sequence 173, Application US/10123262  
 : Publication No. US20030049816A1

## GENERAL INFORMATION:

: APPLICANT: Baker, Kevin P.  
 : APPLICANT: Beresini, Maureen  
 : APPLICANT: DeForge, Laura  
 : APPLICANT: Desnoyers, Luc  
 : APPLICANT: Filvaroff, Ellen  
 : APPLICANT: Gao, Wei-Qiang  
 : APPLICANT: Gerltsen, Mary E.  
 : APPLICANT: Goddard, Audrey  
 : APPLICANT: Godowski, Paul J.  
 : APPLICANT: Gurney, Austin L.  
 : APPLICANT: Sherwood, Steven  
 : APPLICANT: Smith, Victoria  
 : APPLICANT: Stewart, Timothy A.  
 : APPLICANT: Tumas, Daniel  
 : APPLICANT: Watanabe, Colin K  
 : APPLICANT: Wood, William  
 : APPLICANT: Zhang, Zemin  
 : TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 : FILE REFERENCE: P330R1C38  
 : CURRENT APPLICATION NUMBER: US/10/123,262  
 : CURRENT FILING DATE: 2002-04-15  
 : Prior Application removed - See File Wrapper or Palm  
 : NUMBER OF SEQ ID NOS: 550  
 : SEQ ID NO 173  
 : LENGTH: 1210  
 : TYPE: DNA  
 : ORGANISM: Homo Sapien  
 : US-10-123-262-173

## Alignment Scores:

Pred. No.: 1.01e-110 Length: 1210  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3  
 Query Match: 97.81% Indels: 0  
 DB: Gaps: 0

US-10-082-502-19 (1-182) x US-10-123-262-173 (1-1210)

QY 1 MetlySGlyTPGlyTPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTPAla 20  
 Db 543 ATGAAGGCTGGGTTGGTGGCCCTCTCTGGGGGCCCTGCTGGAAACCGCTGGGCT 602  
 QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTP 40  
 Db 603 CGAGAGGCCAGGATCTCCACTGTGAGCAATGCAAGGCGCTGCTGGATGAATGAGATGG 662  
 QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 Db 663 GAATTTGCCAGGTGACCCCAAGAACCATTCAGATGGGATCTTCCGATCAATCCA 722  
 QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
 Db 723 GATGGCAGCCAGTCAAGTGTGAGAGTGCTTATGCCCGCTCAGAGGCCACCTCACAGAG 782  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100

Db 783 CTGCTGAGAGATATATGTGACCGGATGAAGAGATATGGGGAACAGATTCCTCCACC 842  
 QY 101 HisArgLysAsnTyrValArgValValSerArgAsnGlyGluSerSerGluLeuAspLeu 120  
 Db 843 CATCGCAAGAACTACGTACGTGTAGTGGCCGGAATGAGAAATCCAGTAACCTGACCTA 902  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 Db 903 CAAGGCATCCGAATCGACTCAGATTTTACCGGCCACCTCAAGTTTGGCTGTGAGAGCAT 962  
 QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
 Db 963 GTGAGAGATACGAGATGAATCATTCATTCATTTCCCGAGAGGCTGACAAATGTTAA 1022  
 QY 161 AsplysleuCySerlyArgThrAspleuCyAspHisAlaLeuHisArgSerHisAsp 180  
 Db 1023 GACAACTTTGCAGTACGAGACAGATCTTTGTGACCATGCCCTGCACATATCGATGAT 1082  
 QY 181 Glueu 182  
 Db 1083 GAGCTA 1088

Search completed: March 31, 2003, 11:59:08  
 Job time : 59.0233 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:33:58 ; Search time 39.6802 Seconds

(without alignments)  
1406.624 Million cell updates/sec

Title: US-10-082-502-19

Perfect score: 959  
Sequence: 1 MKGCMALLLGLVLTGTAWA.....LCSKRTDLCDAHLSHDEL 182

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame\_p2n.model -DEV-xlp  
-Q-/cgn2\_1/USPTO\_spool/US10082502/runat\_24032003.135100.6616/app.query.fasta.1.1308  
-DB-Issued\_Patents\_NA -QFMT-fastest -SUFFIX-rn1 -MINMATCH-0.1 -LOOPCL-0  
-LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-Blosum62 -TRANS-human40.cdl  
-LIST-45 -DOCALLIGN-200 -THR SCORE-pct -THR MAX-100 -THR MIN-0 -ALIGN-15  
-MODE-LOCAL -OUTFMT-pio -NOR-ext -HEAPSIZ-500 -MINLEN-0 -MAXLEN-2000000000  
-USER-US10082502.ecgn1.1.131-@runat\_24032003.135100.6616 -NCP-6 -ICPU-3  
-NO\_XLPHY -NO\_MAP -LARGEQUERY -NEG\_SCORES-0 -WAIT -LONGLOG -DEV\_TIMEOUT-120  
-NARN\_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6 -FGAPEXT-7  
-XGAPOP-10 -XGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/plodata/1/lna/5A.COMB.seq:\*  
2: /cgn2\_6/plodata/1/lna/5B.COMB.seq:\*  
3: /cgn2\_6/plodata/1/lna/6A.COMB.seq:\*  
4: /cgn2\_6/plodata/1/lna/6B.COMB.seq:\*  
5: /cgn2\_6/plodata/1/lna/6C.COMB.seq:\*  
6: /cgn2\_6/plodata/1/lna/6D.COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	136.5	14.2	744	4	US-09-163-285-3
2	136.5	14.2	1512	4	US-09-163-285-1
3	100	10.4	915	4	US-09-724-864-14
4	87	9.1	5181	1	US-08-257-073-10
5	84	8.8	1530	4	US-09-131-831B-2
6	84	8.8	1554	3	US-08-809-999D-1
7	84	8.8	1554	4	US-09-069-637-1
8	84	8.8	1554	4	US-09-322-360-1
9	84	8.8	1554	4	US-09-131-831B-1
10	84	8.8	1555	2	US-08-316-231B-1
11	84	8.8	2148	3	US-08-809-999D-2
12	84	8.8	2148	3	US-09-069-637-2

13	84	8.8	2148	4	US-09-322-360-2	Sequence 2, Appl1
14	82.5	8.6	4403765	4	US-09-103-840A-2	Sequence 2, Appl1
15	82.5	8.6	4411529	4	US-09-103-840A-1	Sequence 1, Appl1
16	82	8.6	6519	1	US-08-588-985-1	Sequence 1, Appl1
17	82	8.6	6519	1	US-08-971-988-1	Sequence 62, Appl1
18	81.5	8.5	61663	4	US-09-453-702B-62	Sequence 38, Appl1
19	79.5	8.3	5661	4	US-08-838-105-2	Sequence 38, Appl1
20	78	8.1	4131	1	US-07-872-644-38	Sequence 38, Appl1
21	78	8.1	4131	1	US-08-297-494-38	Sequence 38, Appl1
22	78	8.1	4131	1	US-08-297-510-38	Sequence 38, Appl1
23	78	8.1	4131	1	US-08-479-533-38	Sequence 38, Appl1
24	78	8.1	4131	1	US-08-455-526-38	Sequence 38, Appl1
25	78	8.1	4131	1	US-08-455-526-38	Sequence 38, Appl1
26	78	8.1	4131	3	PCT-US92-03222-38	Sequence 38, Appl1
27	78	8.1	4131	5	PCT-US92-03222-38	Sequence 38, Appl1
28	77.5	8.1	1924	4	US-08-961-083-159	Sequence 159, App
29	76.5	8.0	3752	4	US-08-961-527-208	Sequence 208, App
30	76	7.9	833	2	US-08-837-029-1	Sequence 1, Appl1
31	76	7.9	3789	1	US-07-872-644-42	Sequence 42, Appl1
32	76	7.9	3789	1	US-08-297-494-42	Sequence 42, Appl1
33	76	7.9	3789	1	US-08-297-510-42	Sequence 42, Appl1
34	76	7.9	3789	1	US-08-479-533-42	Sequence 42, Appl1
35	76	7.9	3789	1	US-08-455-526-42	Sequence 42, Appl1
36	76	7.9	3789	1	US-08-455-526-42	Sequence 42, Appl1
37	76	7.9	3789	3	US-09-139-491-42	Sequence 42, Appl1
38	76	7.9	3789	5	PCT-US92-03222-42	Sequence 103, App
39	76	7.9	5558	4	US-08-961-527-103	Sequence 3, Appl1
40	75.5	7.9	1216	4	US-09-071-035-3	Sequence 1, Appl1
41	75.5	7.9	1347	4	US-09-071-035-1	Sequence 1, Appl1
42	74.5	7.8	4371	1	US-08-803-972-1	Sequence 1, Appl1
43	74.5	7.8	4371	1	US-08-803-972-1	Sequence 9, Appl1
44	74	7.7	918	3	US-08-937-271-9	Sequence 20, Appl1
45	74	7.7	1349	1	US-07-951-715A-20	

## ALIGNMENTS

RESULT 1  
US-09-163-285-3  
; Sequence 3, Application US/09163285  
; Patent No. 6204013  
; GENERAL INFORMATION:  
; APPLICANT: Rhododoust, Mehran  
; TITLE OF INVENTION: NOVEL MSP-5 PROTEIN AND NUCLEIC ACID MOLECULES  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/163,285  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/090,398  
; FILING DATE: June 24, 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragoras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: MNT-049  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 3:

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 744 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..744
: US-09-163-285-3
:
: Alignment Scores:
: Pred. No.: 6,43e-08 Length: 744
: Score: 136.50 Matches: 49
: Percent Similarity: 44.68% Conservative: 35
: Best local Similarity: 26.06% Mismatches: 61
: Query Match: 14.23% Indels: 43
: DB: Gaps: 10
:
US-10-082-502-19 (1-182) x US-09-163-285-3 (1-744)
OY 11 LeuGIyValLeuLau-----GlyThrIatRpaLa-----Arg 21
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 16 TTGGCAATATTGCTTTTCCTTTTGTGGCCCTGCACAGAGCCTTGCGATGTTGAAG 75
OY 22 ArgSerInaSp-----LeuHIScysGlYalAcYSArgAlaLeuVal 35
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 76 GAGAGAGCAGTGAACACAGAAGCGCTGCCCAAAATGCGAAGGTGTAAAGCTGTGACC 135
OY 36 AspGIueugIuTPgiUilLealargValaSP---ProLYsluSthrIIegImetgLy 54
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136 ACAGAGCTACAGCGCGAGCATGATGCGCACCGCTGCATCTCGAGAGGTGTGACTGGG 195
OY 55 SerPhearGIleaSnProAspgILSerGIuSeRValValgluValProTYRaLaArgSer 74
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 196 CAGGTGCTG-----GATACAGCGACAGAGAAAGACACAGCGCTTACAGCGTTTCA 248
OY 75 GluAlaHISleuThrGIuleuLeuGIuValIcYSasparmetlYsgIutryGLyu 94
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 247 GAGACAAAGCGTAGAAGCGCCTTAGAGCAATTTATGTAGCGGAGCTGACTG----- 300
OY 95 GluIlleaspProSerThrHIS----ArgLYsaNtYrValaTrgValalSerArgasn 112
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 -----AGTGTTACAGCTGAGAGCGCAAGGCGTCACAGATATGCCAAGCGTCAG 348
OY 113 GlyGIuSer-----SerGIuleuaspLeuGIuIlglyIleargIIleasPSerasp 128
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 349 AGTCAGACCATGCCAACACTGMAAGGCGCTACTGCAGAGGGGGTGAAGGTGGATCTGGCG 408
OY 129 IleSer-----GlyThrLeuLYspHealaCYs 137
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 409 ATCCCTCTGGAGGCTTTGGGATGAGACCAGCGTSGAGGTGCACATACCTCCAMAGACAAGTGT 466
OY 138 GluseriLevalIGluGIuTyTrGIuaSpbluLeuIIlegu--PhePheSerargIuaLa 156
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 469 GAGACCAATGTGGAGGAGCTTTGAAGACATTTGTGGAGACTGTGTACTTCACACATCAGAGAG 528
OY 157 AspaNvaNValysAspLYlaCYs 164
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 529 CAGCCCCCTACAAAATTTTCTCTGT 552
:
RESULT 2
US-09-163-285-1
: Sequence 1, Application US/09163285
: Patent No. 6204013
: GENERAL INFORMATION:
: APPLICANT: Khodadoust, Mehran
: TITLE OF INVENTION: NOVEL MSP-5 PROTEIN AND NUCLEIC ACID MOLECULES
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD, LLP
: STREET: 28 State Street

```

```

      CITY: Boston
      STATE: Massachusetts
      COUNTRY: USA
      ZIP: 02109

      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/163,285
      FILING DATE:
      CLASSIFICATION:
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: 60/090,398
      FILING DATE: June 24, 1998
      ATTORNEY/AGENT INFORMATION:
      NAME: Mandragouras, Amy E.
      REGISTRATION NUMBER: 36,207
      REFERENCE/DOCKET NUMBER: MNI-049
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (617)227-7400
      TELEFAX: (617)742-4214
      INFORMATION FOR SEQ ID NO: 1:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 1512 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
      MOLECULE TYPE: cDNA
      FEATURE:
      NAME/KEY: CDS
      LOCATION: 125..868
      US-09-163-285-1

      Alignment Scores:
      Pred. No.: 1,79e-07 length: 1512
      Score: 136.50 Matches: 49
      Percent Similarity: 44.68% Conservative: 35
      Best Local Similarity: 26.06% Mismatches: 61
      Query Match: 14.23% Indels: 43
      DB: Caps: 10

      US-10-082-502-19 (1-182) x US-09-163-285-1 (1-1512)
      Oy 11 LeuGlyValLeuLeu-----GlyThrAlaTrpAla-----Arg 21
      |||||||:||||| |||||||
      Db 140 TTGGGAATATGCTTTCCCTTTTGTGGCCGTGCAGCAGCGCTGGGTGGAGATGTGAG 199
      |||||||:|||||

      Oy 22 ArgSerGlnAsp-----LeuHisCysGlyAlaCysArgAlaLeuVal 35
      |||||:||||| |||||

      Db 200 GAGGAGGACGATGACACAGAACGCTTGCCACGAATGCCAAGTGTTAACTGCTAGC 258
      |||||:|||||

      Oy 36 AspGluLeuGluTrpGluIleAlaArgValAsp--ProIuStyTrhIleGlnMetGly 54
      |||||||:|||||:|||||
      Db 260 ACAGAGCTACAGCGGCAAGCTGAGTCGCACCGGTGATCTCGAGAGGTGTGGAGCTGGGG 319
      |||||||:|||||:|||||

      Oy 55 SerPheArgIleAsnProAspGlySerGlnSerValValGluValProTyrAlaArgSer 74
      |||||:|||||:|||||
      Db 320 CAGGTGCTG-----GATACAGCGCAAGAGAGACACGCTTCACAGCGTTTCA 370
      |||||:|||||

      Oy 75 GluAlaHisLeuTrhGluLeuLeuGluGluValCysAspArgMetLysGluTyrGlyGlu 94
      |||||:|||||:|||||
      Db 371 GAGACAAAGCGCTGGAAGAGCGCTTGAAGAATTATGATAGCGGAGATCCCGAGACTAT 424
      |||||:|||||:|||||

      Oy 95 GlnIleAspProSerTrhHis-----ArgLysAsnTyrValArgValValSerArgAsn 112
      |||||:|||||:|||||
      Db 425 -----ACTGTTCAACGCTGACGCGCAAGCGCTCATGAGATATGCCAAGGTCAG 472
      |||||:|||||:|||||

      Oy 113 GlyGluSer-----SerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 128
      |||||:|||||:|||||
      Db 473 AGTGAACCATGCGCAACACTGAAGAGCGCTGATGTCAGAAAGGGGCTGAAGTGTGATCTGGGG 532
      |||||:|||||:|||||
  
```



QY 129 ILeSer-----GlyThreulysPhealacys 137  
||| ||||| |||  
Db 533 ATCCCTCGAGCTTTGGGATGAGCCAGCGTGAGGTCTACATCTACCTAAGAGAGCTGT 592  
QY 138 GluSerIleValGluGluTyrGluAspGluLeuIleGlu---PhePheSerArgGluAla 156  
|||||:|||||:|||||: |||: |||: |||: |||  
Db 593 GAGACCATGTTGGAGAGATTTGAGACATTTGGGAGACTGTCTACTTCCACCATGAGAG 652  
QY 157 AspaValIlyAspIlyLeucys 164  
:|||||: |||||  
Db 653 CAGCCCTACAAATTTCTCTGT 676  
RESULT 3  
US-09-724-864-14  
: Sequence 14, Application US/09724864  
: Patent No. 6380362  
: GENERAL INFORMATION:  
: APPLICANT: Watson, James D.  
: APPLICANT: Murison, James G.  
: TITLE OF INVENTION: Polynucleotides, polypeptides expressed  
: FILE REFERENCE: 11000.105001  
: CURRENT APPLICATION NUMBER: US/09/724,864  
: PRIOR FILING DATE: 2000-11-28  
: PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678  
: NUMBER OF SEQ ID NOS: 72  
: SOFTWARE: FastSeq for Windows Version 4.0  
: SEQ ID NO 14  
: LENGTH: 915  
: TYPE: DNA  
: ORGANISM: Mouse  
US-09-724-864-14  
Alignment Scores:  
Pred. No.: 0.00303 Length: 915  
Score: 100.00 Matches: 49  
Percent Similarity: 39.71% Conservative: 32  
Best Local Similarity: 24.02% Mismatches: 73  
Query Match: 10.43% Indels: 50  
DB: 4 Gaps: 10  
US-10-082-502-19 (1-182) x US-09-724-864-14 (1-915)  
QY 7 LeuAlaLeuLeuLeuGly-----ValLeuLeuGlyThrAlaArgSer--- 23  
||| ||||| ||| :|||:|||||: ||| |||  
Db 38 CTGTGTCTACTGTTGGGTGAGGGCTATCTGAGGAGCGCGGGATGAGGTTTCCCTC 97  
QY 24 -----GlnAsp 25  
Db 98 TCGGCTTCGGCTCCACACTGATGATGAGAGAGAGTACTCGGCTCATATCGCGGCTCAC 157  
QY 26 LeuHleGlyAlaGlyAlaGlyAlaLeuValAspGluLeuGluTrrGluIleAlaArgVal 45  
||| ||| |||||:|||||: |||: |||: |||: |||  
Db 158 CTGCGCTGCGCATGCTCGCGGCTGTGGCTTCACAGTGGGAGACGCTTCGCGGAAAGCA 217  
QY 46 AspProLysLysThrIleGlnMetGlySerPheArgIleAsnProAspGlySerGlnSer 65  
: ||| :||| :|||: |||: |||: |||  
Db 218 GAGGCTAAATCTCAC-----ACTGCAGAGCGCAAT---GGA 250  
QY 66 ValValGluValProTyrAlaArgSerGluAlaHisLeuThrGluLeuGluGluVal 85  
: ||| :||| :|||: |||: |||: |||: |||: |||  
Db 251 TTTCAGAGAGCTG-----AGTGAATCCACGATGATGATGCTGCGGACAGAC 298  
QY 86 CysAspArg---MetLysGluTyrGluGluIleAspProSerThrHisArgLysAsn 104  
||| :||| :|||: |||: |||: |||  
Db 299 TCGCTCAGACAGCGAGCTCTATGCA-----CTCATGAGTGAAC 340  
QY 105 TyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArg 124  
: ||| :||| :|||: |||: |||: |||  
Db 341 CAGATGAGCGCTCTCAGCGGCGCAGAGACTTAGCAGAGGGCGCAGAG-----CCAAAG 391  
QY 125 IleAspSerAspIleSerGly-----ThreulysPheAlaCysGluSer 139

Db 392 ATCAGCGTGATGATTTGGGGGTCCTCGGCCCAATAGCTCTCTCCAGAGCTGTTTCCAC 451  
QY 140 IleValGluGluTyr---GlnAspGluLeuIleGluPheSerArgGluAlaAspAsn 158  
: |||: |||: |||: |||: |||: |||: |||: |||  
Db 452 TACCTGGGTGATTTGGAGAGACCATGATCTATGAGCGCTACCGGACGCGCAAGCGCAAT 511  
QY 159 ValLysAspIlyLeuCysSerLysArgThrAspLeuGlyAspHisAlaLeuHisArgSer 178  
: |||: |||: |||: |||: |||: |||  
Db 512 CTGAGAGCGGCTGCTCTGTGGGGGACACCATGAGCGCTGCTCAGAGAGATCTGGCCAG 571  
QY 179 HisAspGluLeu 182  
: |||: |||: |||  
Db 572 AGAGAGAGCTT 583  
RESULT 4  
US-08-257-073-10  
: Sequence 10, Application US/08257073  
: Patent No. 5766597  
: GENERAL INFORMATION:  
: APPLICANT: Paolelli, Enzo  
: APPLICANT: de Talsene, Charles  
: TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE  
: NUMBER OF SEQUENCES: 143  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Curtiss, Morris & Safford, P.C.  
: STREET: 530 Fifth Avenue, 25th Floor  
: CITY: New York  
: STATE: New York  
: COUNTRY: UNITED STATES OF AMERICA  
: ZIP: 10036  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/257,073  
: FILING DATE: 09-JUN-1994  
: CLASSIFICATION: 424  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/075,783  
: FILING DATE: 11-JUN-1993  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 07/852,305  
: FILING DATE: 18-MAR-1992  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 07/672,183  
: FILING DATE: 20-MAR-1991  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Frommer, William S.  
: REGISTRATION NUMBER: 25,506  
: REFERENCE/DOCKET NUMBER: 454310-2570  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (212) 840-3333  
: TELEFAX: (212) 840-0712  
: TELEX: 425066 CURTMS  
: INFORMATION FOR SEQ ID NO: 10:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 5181 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
US-08-257-073-10  
Alignment Scores:  
Pred. No.: 1.53 Length: 5181  
Score: 87.00 Matches: 36  
Percent Similarity: 43.95% Conservative: 33  
Best Local Similarity: 22.93% Mismatches: 58  
Query Match: 9.07% Indels: 30  
DB: 1 Gaps: 7

US-10-082-502-19 (1-182) x US-08-257-073-10 (1-5181)

OY 36 AspGluLeuGluTyrPgiuIleAlaArgValAspProLysylSerThrIleGlnMetGlySer 55  
||| ||||||| ||| : : : ||| : : :  
Db 1192 GATCCACTGGAATTGAAATATTTATTNAGAGAAAAAATAAGTGTACACCT 1251

OY 56 PheArgIleAsnProAspArgGlySerGlnSerValValGluValProTyrAlaArgSerglu 75  
:::||| ||| : : : |||||||  
Db 1252 AATTCACAAGAATCCTACGAATATCGTTCAATACCAAAAATCTCTTAATGGTATT 1311

OY 76 AlaHis-----LeuThrGluLeuLeuGluValCys-----AspArgMetLys 90  
: : : ||||||| : : :  
Db 1312 GTATATCTCTTACCACCTCACCTGATATTCATTAATTCATTACTCTGCAGATATGATAAAAT 1371

OY 91 GluTyrGluGluGlnIleAspProSerThrHisArgLysAspTyrValArGValValSer 110  
||| ||| : : : ||| : : :  
Db 1372 TCATATGCTGTATTAAATGAAATCCGATACTAAAGAAAATTAATGAAAATTAATTAACA 1431

OY 111 ArgAsnGlyglu-----SerSergLeuAspLeuGlnGly 122  
||| ||| : : : ||| : : :  
Db 1432 GATTAATTAAGCAAAGAAAAATATTCATTAAATACATTAAAAACAATTCATTAGAACAA 1491

OY 123 IleArgIleAspSeraSpIleSergLyThrLeuLysPheAlaCysGluSerIleValIgu 142  
: : : ||| : : :  
Db 1492 AAAAAAATTATCATC-----ACAAAAGAACAAATATAAAAAATTACTTGAA 1536

OY 143 GluTyrGlu-----AspGluLeuIleGlu---PhePheSeraArgLyu 155  
: : : ||| : : : ||| : : :  
Db 1537 GATTATGAAAGATCAAAAAAGGATTATGAAAGAAATTCCTGAAAAAATTTTATGAAATGAAA 1596

OY 156 AlaAspAsn-----ValLysAspLysLeuCysSerLysArg 167  
: : : ||| ||| ||| : : : |||  
Db 1597 TTTAATATATATTTTGACAAAGATGCTGATGATTAATAATATTCAGTCGACAGA 1647

RESULT 5  
US-09-131-831B-2  
Sequence 2, Application US/09131831B  
Patent No. 6339149

GENERAL INFORMATION:  
APPLICANT: Coulie, Pierre; Ikeda, Hideyuki; Boon-  
Fallour, Thierry  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules  
Coding For Tumor Rejection Antigen Precursors DAGE and  
Uses Thereof

NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski L.L.P.  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 Inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/131,831B  
FILING DATE: 11-Aug-1998

CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/809,999  
FILING DATE: 9-April-1997  
APPLICATION NUMBER: 08/316,231  
FILING DATE: 30-September-1994

ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6339149man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5386.3

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3100

```

: TELEFAX: (212) 318-3400
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
:     LENGTH: 1530 base pairs
:     TYPE: nucleic acid
:     STRANDEDNESS: double
:     TOPOLOGY: linear
: MOLECULE TYPE: nucleic acid
: FEATURE:
:     NAME/KEY: DAGE (H12)
: SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-131-831B-2

Alignment Scores:
Pred. No.:          0.624          Length:          1530
Score:              84.00          Matches:          40
Percent Similarity: 38.67%         Conservative:    30
Best Local Similarity: 22.10%      Mismatches:     61
Query Match:        8.76%          Indels:         50
DB:                 4              caps:           8

US-10-082-502-19 (1-182) x US-09-131-831B-2 (1-1530)
Oy      1  MetLysGIYTrpGIYTrpLeuAlaLeuLeuLeuGlyValLeuLeuGlyThrAlaIrrpAla 20
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      217  GTGAGGCGCTGGCCCTCTACCTGCCCTCCCTTGGGAGTGGCTGATG----- 261
Oy      21  ArgArgSerGlnAspLeuHisCysGlyValAcysArgAlaLeuValAspGluLeuGlu--- 39
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      262  ---AAGGACAAACATCTTCACCTGGAGACCTTCAAAGCTGTGCTGATGACCTGATGTG 318
Oy      40  -----TrpGluLeuAlaArgValAspProGlyLys 49
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      319  CTCCTTGCCACGAGGAGTTCGCCCCAGGAGGTGGAAACTTCAGTGTGATTAACGGAG 378
Oy      50  -----ThrIleGlnMetGly-----SerPhe 56
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      379  AACCTCATCAGACTCTTGACTGTGATGTGTGTAAGAACAGGCCACTCTGTACTATT 438
Oy      57  -----ArgIleAsnProAspGlySerGlnSer 65
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      439  CCAGAGCCCAAGAGCAGCTCAGCCCATGACAAAGAAAGGAAAGTAAAGTGTGGACACA 498
Oy      66  ValValGluValProTyrAlaArgSerGluAlaHisLeuThrGluLeuLeuGluGlu--- 84
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      499  GAGCGAGAGCAGCCCTTCATTCACAGTAGAGGTGCTCCTAGACCTGTTCCTCAAGAGAGGT 558
Oy      85  ValCysAspArgMetLysGluTyrGlyGluGlnIleAsnProSerThrHisArgLysAsn 104
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      559  GCCTGTGATGATGATTGTTCTCTTAC-----CTCATTAAGAAAGTGAAGCGAAAGAAAT 612
Oy      105  TyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuValAspLeuGlnGlyIleArg 124
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      613  GTACTACGCCCTGTGTAAGAACCTGAAGATTTTTCGATAGCCCATGACGAG----- 663
Oy      125  IleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGluTyr 144
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      664  -----GATATCAAGATGATGATCTGAATAATGTCGACGCTGACCTATTGAAGATT 714
Oy      145  Glu 145
      :  :  :
Db      715  GAA 717

RESULT 6
US-08-809-999D-1
: Sequence 1, Application US/0880999D
: Patent No. 6013765
: GENERAL INFORMATION:
: APPLICANT: Coulie, Pierre; Ikeda, Hideyuki;
: APPLICANT: Boon-Fallou, Thierry
: TITLE OF INVENTION: Isolated Nucleic Acid Molecules
: TITLE OF INVENTION: Coding for Tumor Rejection Antigen Precursors DAGE and
: TITLE OF INVENTION: Uses Thereof

```

NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski, L.L.P.  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,999D  
FILING DATE: 9-April-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/316,231  
FILING DATE: 30-September-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6013765man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5386.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3000  
TELEFAX: (212) 752-5958  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1554 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
FEATURE:  
NAME/KEY: DAGE (5E10)  
US-08-809-999D-1

Alignment Scores:  
Pred. No.: 0.638 Length: 1554  
Score: 84.00 Matches: 40  
Percent Similarity: 38.67% Conservative: 30  
Best Local Similarity: 22.10% Mismatches: 61  
Query Match: 8.76% Indels: 50  
DB: 3 Gaps: 8

US-10-082-502-19 (1-182) x US-08-809-999D-1 (1-1554)

QY 1 MetLysGLYTrpGLYTrpLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
DB 276 GTGCAAGGCTGGCCCTTCACCTGCTCCCTGCGAGTGTGATG----- 320  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGlu--- 39  
DB 321 ---AAGGACACACATCTTCACCTGAGACCTTCAAAAGCTGTGTTGATGAGCTATGATGTC 377  
QY 40 -----TrrGluIleAlaArgValAspProLysLys 49  
DB 378 CTCCTTGGCCAGAGGTTGCCGCCAGAGAGGTGGAAACTTCAGTCTGATTTACGGAAG 437  
QY 50 -----ThrlleGlnMetGly-----Serphe 56  
DB 438 AACTCTCATCAGAGACTTCTGGACTGTATGTTGGAAACAGGCCACGTCTACTCATTTT 497  
QY 57 -----ArgIleAspProAspGlySerGlnSer 65  
DB 498 CCAGAGCCAGAGACAGCTCAGCCCATGACAAAGAGCGAAAGATAGATGTTTGACACACA 557  
QY 66 ValValGluValProGlyTrpAlaArgSerGluValAlaHisLeuThrGluLeuGluGlu--- 84  
DB 558 GAGGACAGACACACCTTTCATTCAGTAGAGAGGTGCTGTGACCTGTCTTCCAGGAAGGT 617  
QY 85 ValCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThrHisArgLysAsn 104

DB 618 GCCTGATGATAATTTGTTCTCTAC-----CTCAATTGAGAAAGTGAAGCCAGAAAAAAT 671  
QY 105 TyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArg 124  
DB 672 GTACTACGCTGCTGCTTAAGAACTCAAGATTTTTCATTCACCCATGACG----- 722  
QY 125 IleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGluTrp 144  
DB 723 -----GATATCAAGATGATCCTCGAATAATGCTGACGTGACCTCTATTGAAGATTG 773  
QY 145 Glu 145  
DB 774 GAA 776

RESULT 7  
US-09-069-637-1  
Sequence 1, Application US/09069637  
Patent No. 6022692  
GENERAL INFORMATION:  
APPLICANT: Coullie, Pierre; Ikeda, Hideyuki;  
APPLICANT: Boon-Falleur, Thierry  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules  
TITLE OF INVENTION: Coding For Tumor Rejection Antigen Precursors DAGE and Uses  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felte & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/069,637  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,999  
FILING DATE: 9-April-1997  
APPLICATION NUMBER: 08/316,231  
FILING DATE: 30-September-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6022692man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5386.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1554 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
FEATURE:  
NAME/KEY: DAGE (5E10)  
US-09-069-637-1

Alignment Scores:  
Pred. No.: 0.638 Length: 1554  
Score: 84.00 Matches: 40  
Percent Similarity: 38.67% Conservative: 30  
Best Local Similarity: 22.10% Mismatches: 61  
Query Match: 8.76% Indels: 50  
DB: 3 Gaps: 8

US-10-082-502-19 (1-182) x US-09-069-637-1 (1-1554)



TITLE OF INVENTION: Isolated Nucleic Acid Molecules  
Coding For Tumor Rejection Antigen Precursors DAGE and  
Uses Thereof  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski L.L.P.  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/131,831B  
FILING DATE: 11-Aug-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/809,999  
FILING DATE: 9-Apr-11-1997  
APPLICATION NUMBER: 08/316,231  
FILING DATE: 30-September-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6339149man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5386.3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3100  
TELEFAX: (212) 318-3400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1554 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
FEATURE:  
NAME/KEY: DAGE (5E10)  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-131-831B-1  
Alignment Scores:  
Pred. No.: 0.638 Length: 1554  
Score: 84.00 Matches: 40  
Percent Similarity: 38.67% Conservative: 30  
Best Local Similarity: 22.10% Mismatches: 61  
Query Match: 8.76% Indels: 50  
Gaps: 8  
US-10-082-502-19 (1-182) x US-09-131-831B-1 (1-1554)  
QY 1 MetLygGlyTTPolYyTTPleuAlaleuLeuLeuGlyYValleuGlyYThraAlaTTPala 20  
DB 276 GTGACAGCGCTGGCCCTGACCTGCTCCCTGCGAGTGTGATG----- 320  
QY 21 ArgARSerGlnaSprLeuHiaCysGlyAlaCysArgAlaLeuValaSprGluLeuGlu--- 39  
DB 321 ---AAGGACACATCTTCACTCGAGACCTTAAAGCTGTCTTGATGACCTTGATGTG 377  
QY 40 -----TTPGluIleAlaArgValaSprProLyLyS 49  
DB 378 CTCCTTCCAGAGAGGTGCGCCAGAGAGGTGAACCTCAAGTGTGATTTACGGAAG 437  
QY 50 -----ThrlleGlnMeGly-----SerPhe 56  
DB 438 AACCTCATCAGAGACTTGTGGACTGTATGGTGTGAACAGGCGCACTGTACTATT 497  
QY 57 -----ArgTleasnProAspGlySerGlnSer 65  
DB 498 CCAGAGCCAGAGACGACTGACCCCATGACAAAGAGGAAAGTAAGTGTGTTGAGCACA 557

QY 66 ValValGluValProTyrAlaArgSerGluAlaIleuThrlGluLeuGluGlu---- 84  
DB 558 GAGGACAGACGCGCTTCATTCAGTAGAGTGTGAGACSTGTCTCCAGGAAGGT 617  
QY 85 ValCysAspArgMetLySgluTyrGlyGluGlnIleAspProSerThrIleAspGlySaa 104  
DB 618 GCCTGTGATGAATGTGTCTCTAC-----CTCATTTGAGAAAGTGAAGCGAAAGAAAT 671  
QY 105 TyrValArgValValSerArgaenGlyGluSerSerGluLeuAspLeuGlnGlyIleArg 124  
DB 672 GTACTACGCTGTGCTGTAGAAGCTGAAGATTTTGCATGTCCATGCCAG----- 722  
QY 125 IleAspSerAspIleSerGlyThrlLeuLySPhleAlaCysGluSerIleValGluGluTyr 144  
DB 723 -----GATATCAAGATGATCTGAAAGATGTCAGCTGTGATCTGTATGAAGATTG 773  
QY 145 Glu 145  
DB 774 GAA 776  
RESULT 10  
US-08-316-231B-1  
Sequence 1, Application US/08316231B  
Patent No. 5830753  
GENERAL INFORMATION:  
APPLICANT: Coullie, Pierre; Ikeda, Hideyuki;  
APPLICANT: Boon-Falleur, Thierry  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules  
TITLE OF INVENTION: Coding For Tumor Rejection Antigen Precursors DAGE and  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felte & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/316,231B  
FILING DATE: 30-September-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5830753man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5386  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1555 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
FEATURE:  
NAME/KEY: DAGE  
OTHER INFORMATION: Xaa is Arg when V is C or A or  
OTHER INFORMATION: Gly when V is G  
US-08-316-231B-1  
Alignment Scores:  
Pred. No.: 0.639 Length: 1555  
Score: 84.00 Matches: 40  
Percent Similarity: 38.67% Conservative: 30  
Best Local Similarity: 22.10% Mismatches: 61

```

Query Match:      8.76%      Indels:      50
DB:               2          Gaps:         8

US-10-082-502-19 (1-182) x US-08-316-231b-1 (1-1555)

OY      1 MetLysGlyTPrGlyTPrLeuAlaLeuLeuGlyValLeuGlyThrAlaTPrAla 20
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      277 GTGAGGCGCTGGCCCTTACCTGCTCCCTGTGGAGTGTGATG----- 321

OY      21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGlu--- 39
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      322 ---AAGGACAAACATCTTCACTGAGACCTTCAAGCTGCTGATGAGCTTATG 378

OY      40 -----TPrGluIleAlaArgValAspProLys 49
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      379 CTCCTGGCCAGAGTTCCGCCAGAGAGTGAACCTTCAAGTCTGATTAAGGAA 438

OY      50 -----ThrIleGlnMetGly-----SerPhe 56
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      439 AACTCTATCAGACTTCTGAGTGTATGCTGTGAAACAGGCCAGCTGTACTCATTT 498

OY      57 -----ArgIleAsnProAspGlySerGlnSer 65
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      499 CCAGAGCCAGAAAGCAGCTCAGCCCATGACAAAGAAAGTAAAGTGTGAGCACA 558

OY      66 ValValGluValProTyrAlaArgSerGluAlaHisLeuThrGluLeuGluGlu--- 84
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      559 GAGCAGACGAGCCCTTCATTCAGTAGAGCTGCTGTAAGACCTGTCTTCCAGAGAGCT 618

OY      85 ValCysAspArgMetLysGlyTyrGlyGluGlnIleAsnProSerThrHisArgLysAsn 104
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      619 GCGCTGTATGAATGTCTCTAC-----CTCATTTGAAAGTGAAGCAGAAAGAAAT 672

OY      105 TyrValArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeuGlnGlyLea 124
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      673 GTACTAGCCCTGTGTAAGAGCTGAAGATTTTTCATGAGCCCATG----- 723

OY      125 IleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlySerIleValGluGluTyr 144
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      724 -----GATATCAGATGATCTGTGAATAATGGTGCAGCTGAGCTATTTGAAGATTGG 774

OY      145 Glu 145
      : : :
DB      775 GAA 777

RESULT 11
; Sequence 2, Application US/08809999D
; Patent No. 6013765
; GENERAL INFORMATION:
; APPLICANT: Coullie, Pierre; Ikeda, Hideyuki;
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules
; TITLE OF INVENTION: Coding For Tumor Rejection Antigen Precursors DAGE and
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,999D
; FILING DATE: 9-April-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: 08/316,231
FILING DATE: 30-September-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6013765man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5386.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2148 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid
FEATURE:
NAME/KEY: DAGE (H12)
US-08-809-999D-2

Alignment Scores:
Pred. NO.:      1.02      Length:      2148
Score:          84.00     Matches:      40
Percent Similarity: 38.67%   Conservative: 30
Best Local Similarity: 22.10%   Mismatches: 61
Query Match:      8.76%     Indels:      50
DB:               3          Gaps:         8

US-10-082-502-19 (1-182) x US-08-809-999D-2 (1-2148)

OY      1 MetLysGlyTPrGlyTPrLeuAlaLeuLeuGlyValLeuGlyThrAlaTPrAla 20
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      452 GTGAGGCGCTGGCCCTTACCTGCTCCCTGTGGAGTGTGATG----- 496

OY      21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGlu--- 39
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      497 ---AAGGACAAACATCTTCACTGAGACCTTCAAGCTGTGCTTGAAGCTTATG 553

OY      40 -----TPrGluIleAlaArgValAspProLys 49
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      554 CTCCTGGCCAGAGTTCCGCCAGAGAGTGAACCTTCAAGTCTGATTAAGGAA 613

OY      50 -----ThrIleGlnMetGly-----SerPhe 56
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      614 AACTCTATCAGACTTCTGAGTGTATGCTGTGAAACAGGCCAGCTGTACTCATTT 673

OY      57 -----ArgIleAsnProAspGlySerGlnSer 65
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      674 CCAGAGCCAGAAAGCAGCTCAGCCCATGACAAAGAAAGTAAAGTGTGAGCACA 733

OY      66 ValValGluValProTyrAlaArgSerGluAlaHisLeuThrGluLeuGluGlu--- 84
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      734 GAGCAGACGAGCCCTTCATTCAGTAGAGTGTGTCGTAAGACCTGTCTCAAGAGAGT 793

OY      85 ValCysAspArgMetLysGlyTyrGlyGluGlnIleAsnProSerThrHisArgLysAsn 104
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      794 GCGTGTATGAATGTCTCTAC-----CTCATTTGAAAGTGAAGCAGAAAGAAAT 847

OY      105 TyrValArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeuGlnGlyLea 124
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      848 GTACTAGCCCTGTGTAAGAGCTGAAGATTTTTCATGAGCCCATGAG----- 898

OY      125 IleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlySerIleValGluGluTyr 144
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      899 -----GATATCAGATGATCTGTGAATAATGGTGCAGCTGAGCTATTTGAAGATTGG 949

RESULT 12
; Sequence 2, Application US/09069637

```

Alignment Scores:			
Pred. No.:	1.02	Length:	2148
Score:	84.00	Matches:	40
Percent Similarity:	38.67%	Conservative:	31
Best Local Similarity:	22.10%	Mismatches:	60
Query Match:	8.76%	Indels:	50
DB:	3	Gaps:	8

  

US-10-082-502-19 (1-182) x US-09-069-637-2 (1-2148)
QY 1 MetlySGlyTPRGlyTPRLeuAlaLeuLeuLeuGlyValLeuLeuGlyThrAlaTPRAla 20
Db 452 GTGCAGGCGCTGCCTTACGCTGCTCCCTCGGAGTGTGTGANG----- 496
QY 21 ArgArgSerGlnAAsPLeuHnHScysGlyAlaLeuArgAlaLeuValAAsPLeuLeuGlu-- 39
Db 497 ---AAGGACAAACATCTTCAACCTGAGACGCTTCAAGCGTGGCTGTATGTGACCTGATGTG 553
QY 40 -----TPRLeuAlaAlaArgValAAsPProlyls 49
Db 554 CTTCCTTGCCAGAGAGGTTGCGGCCAGAGAGTGAACACTCAAGTGTGTGATTTACGGAG 613
QY 50 -----ThrlleGlnMetly-----SerPhe 56
Db 614 AACTGTCTATCATAGGACTTCTGTGACTGTATGTCTGTGAAACAAGGCCAAGCTGTACTCATTT 673
QY 57 -----ArgIleAsnPProAAsPLeuSerGlnSer 65

```

1      RESULT 13
2      US-09-322-360-2
3      : Sequence 2, Application US/09322360
4      : Patent No. 6297050
5      : GENERAL INFORMATION:
6      : APPLICANT: Coulle, Pierre; Ikeda, Hideyuki;
7      : APPLICANT: Boon-Falleur, Thierry
8      : TITLE OF INVENTION: Isolated Nucleic Acid Molecules
9      : TITLE OF INVENTION: Coding For Tumor Rejection Antigen Precursors DAGE and
10     : TITLE OF INVENTION: Uses Thereof
11     : NUMBER OF SEQUENCES: 18
12     : CORRESPONDENCE ADDRESSES:
13     : ADDRESSEE: Fulbright & Jaworski, L.L.P.
14     : STREET: 666 Fifth Avenue
15     : CITY: New York City
16     : STATE: New York
17     : COUNTRY: USA
18     : ZIP: 10103
19     : COMPUTER READABLE FORM:
20     : MEDIUM TYPE: Diskette, 3.5 Inch, 1.44 Mb storage
21     : COMPUTER: IBM PS/2
22     : OPERATING SYSTEM: PC-DOS
23     : SOFTWARE: Wordperfect
24     : CURRENT APPLICATION DATA:
25     : APPLICATION NUMBER: US/09/322,360
26     : FILING DATE:
27     : CLASSIFICATION:
28     : PRIOR APPLICATION DATA:
29     : APPLICATION NUMBER: 08/809,999
30     : FILING DATE:
31     : ATTORNEY/AGENT INFORMATION:
32     : NAME: Hanson, No. 6297050man D.
33     : REGISTRATION NUMBER: 30,946
34     : REFERENCE/DOCKET NUMBER: LUD 5386.1
35     : TELECOMMUNICATION INFORMATION:
36     : TELEPHONE: (212) 318-3000
37     : TELEFAX: (212) 752-5958
38     : INFORMATION FOR SEQ ID NO: 2:
39     : SEQUENCE CHARACTERISTICS:
40     : LENGTH: 2148 base pairs
41     : TYPE: nucleic acid
42     : STRANDEDNESS: double
43     : TOPOLOGY: linear
44     : MOLECULE TYPE: nucleic acid
45     : FEATURE:
46     : NAME/KEY: DAGE (H12)
47     : US-09-322-360-2
48
49 Alignment Scores:
50     Fred. No.: 1.02
51     Length: 2148

```

Alignment Scores:		
Pred. NO.:	8.9e+04	Length: 4403765
Score:	82.50	Matches: 44
Percent Similarity:	40.41%	Conservative: 34
Best Local Similarity:	22.80%	Mismatches: 64
Query Match:	8.60%	Indels: 51
DB:	4	Gaps: 12

US-10-082-502-19 (1-182) x US-09-103-840A-2 (1-4403765)

```

Oy      2  LysGIYTRGILTYRPLEALALEUHEULCIGVALLEULEUGILYTThRLaLARG 21
          |||||
Db 4382280 AAGGGCTGGGGGTGGGTGGGCCCGCCSTATTATGAGGAATGC-----TGGCCGAC 4382233
Oy      22  --ArgSerGIAnSPLeUNInIScYSGLValASuARgAlALeUALSPdLUHeUGITPr 40
          |||
Db 4382232 GGGGATTCGAAAGAGTAAGTACGAGCGGCGCT-----GTGGCGTG 4382194
Oy      41  -----GIIPLLaLRgVALAsPRoLYLvsThIle---GLmEtcly 54
          ::::
Db 4382193 CGSAGCGCGGGGCAACGATTCGGCGCTGACATTCAGATCAGACGCGCGGGCGCATGGCG 4382134
Oy      55  SerPheAdgLIlaenPRoARpJLserLINbeRValALGIvualPRoTYrALnArGer 74
          |||
Db 4382133 GTATTATGGCCCCAACAGACTCCCGAGCGAGGGGTGATTCAGATCGCGGTTCGACGCG 4382074
Oy      75  GIUALnISLeUTrGILUeULeu-----GIUGIVuIALSuARgMeTLySLU 91
          |||
Db 4382073 TACGCCAGCACACACCGCTGCTCGGGSCAAATGCCACACAGCTGGCGGCCCATAGACC 4382014
Oy      92  TyrgLgLuGINlLeARProSeThRIaLRgVaaNTYrVALrVALseARg 111
          |||
Db 4382013 TATGCCCGCGCATGCAC--CGCGTGCAT-----4381987
Oy      112 AsnGIUsErSeEGILUeULAsPRUCInELYTleArGILIeARSeARPIleserLy 131
          :::::
Db 4381986 -----GCGGGGTCTCGGATTTGTGGC---GCAATCTCGATCCGCTGACCGGG 4381939
Oy      132 ThRIeULyPRhAIaCySGIsERtLeVALGIUGITyTrGIUALSPdLUe----- 148
          ::|||
Db 4381938 ---ATCAAAGAAGTGTGGAGTTTCTACAGCACAGACGAAGACAGATCSAGCGCATC 4381882
Oy      149  -----TleGIUPhESeARgLIuIAARPAInVALIySAr 161
          ::::
Db 4381881 GCCCATGACATCGCGGTGTGTGACACATTCTACGCGGGAAGN-----GAC 4381834
Oy      162 LysLeuCysSErLyARgThRAspReucYsaRNhIALa 174
          |||
Db 4381833 GCCTTGGCTCGGAGATCACCGCGGTGTGTGCACGCC 4381795

RESULT 15
US-09-103-840A-1/c
: Sequence 1, Application US/09103840A
: Patent No. 6294328
: GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: FILE REFERENCE: 24366-2007.00
: CURRENT APPLICATION NUMBER: US/09/103,840A
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 1
: LENGTH: 4411529
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: OTHER INFORMATION: H37rv
US-09-103-840A-1

```







GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:08:03 ; Search time 142.849 Seconds

(without alignments)  
2869.213 Million cell updates/sec

Title: US-10-082-502-19

Perfect score: 959

Sequence: 1 MKGMGMLALLGLVLTGTA.....LCSTRDLCDAHARSHDEL 182

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/gen2.1/USPFO.spool/US10082502/rnatc\_24032003\_135058.6538/app\_query.fasta\_1.1308  
-DB=N\_Geneseq\_101002 -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsumb2 -TRANS=human40.cdl  
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10082502 -GCN=1.1741 -runat\_24032003\_135058.6538 -KCP=6 -ICFO=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TMEOUT=120  
-WARN\_TMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database: N\_Geneseq\_101002.\*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	959	100.0	1069	20	AAK06970
2	959	100.0	1069	21	AAZ08294
3	938	97.8	546	21	AAZ38326
4	938	97.8	806	20	AAK06969
5	938	97.8	806	21	AAZ08293
6	938	97.8	814	21	AAZ38327
7	938	97.8	814	24	ABK52765
8	938	97.8	814	24	ABK52765
9	938	97.8	814	24	ABK52765
10	938	97.8	814	24	ABK52765
11	938	97.8	814	24	ABK52765
12	931	97.1	814	24	ABK52765
13	896	93.4	657	24	ABK52765
14	889	92.7	657	24	ABK52765
15	856.5	89.3	532	20	AAK97837
16	380	39.6	484	22	AAK43525
17	380	39.6	558	22	AAK25628
18	380	39.6	558	22	AAK25628
19	380	39.6	558	22	AAK25628
20	361	37.6	1160	22	AAK25628
21	361	37.6	649	20	AAK97883
22	361	37.6	649	20	AAK97883
23	361	37.6	649	20	AAK97883
24	358	37.3	718	20	AAK208284
25	357	37.2	792	20	AAK208284
26	357	37.2	415	21	AAK06971
27	347	36.2	1085	20	AAK06971
28	308	32.1	2532	22	AAK07341
29	308	32.1	2532	22	AAK07341
30	308	32.1	2532	22	AAK07341
31	308	32.1	2724	22	AAK07340
32	262	27.3	564	24	AAK32767
33	212.5	22.2	816	23	ABK04671
34	180	18.8	352	22	ABK04671
35	157	16.4	786	21	ABK04671
36	145	15.1	750	21	ABK04671
37	145	15.1	750	21	ABK04671
38	145	15.1	750	21	ABK04671
39	140	14.6	3110	23	ABL21470
40	136.5	14.2	744	21	AAZ49678
41	136.5	14.2	746	22	AAZ49678
42	136.5	14.2	746	22	AAZ49678
43	136.5	14.2	746	22	AAZ49678
44	136.5	14.2	746	22	AAZ49678
45	136.5	14.2	746	22	AAZ49678

## ALIGNMENTS

RESULT 1  
AAK06970  
AAK06970 standard; cDNA; 1069 BP.

AAK06970: 10-MAY-1999 (first entry)

Mouse secretory peptide-9 (Zs199) orthologue cDNA.

Secretory peptide-9: Zs199; orthologue; mouse; tumour marker;

cancer; therapy; diagnosis; growth enhancer; ss.

XX Mus sp.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Location/Qualifiers  
358..906  
/\*tag- a

```

FT      sig_peptide      358..417
FT      /*tag= b
FT      mat_peptide      418..903
FT      /*tag= c
XX
XX      MO9901554-A1.
XX
XX      14-JAN-1999.
XX
XX      02-JUL-1998;      98WO-US13859.
XX
XX      17-JUN-1998;      98US-0099005.
XX      03-JUL-1997;      97US-0051704.
XX      03-JUL-1997;      97US-0888088.
XX      19-MAY-1998;      98US-0081338.
XX      19-MAY-1998;      98US-0085983.
XX      17-JUN-1998;      98US-0089899.
XX
XX      (ZYMO ) ZYMOGENETICS INC.
XX
XX      Jaspers SR, Jellinek LJ, Sheppard PO, Whitmore TE;
XX
XX      WPI; 1999-106055/09.
XX      P-PSDB; AAM88476.
XX
XX      New mammalian secretory peptide-9 (Zsig9) - used as a growth
XX      enhancer for placenta, liver and heart, and as an indicator of
XX      cancer
XX
XX      Claim 2; Page 75-77; 85pp; English.
XX
XX      This CDNA clone encodes novel mouse secretory peptide-9, or Zsig9
XX      (see AAM88476), an orthologue of novel human Zsig9 (see AAM88469).
XX      Human Zsig9 is overexpressed in a number of tumours including
XX      brain, liver, lung, oesophageal, stomach, colon, rectal, thyroid
XX      and lymphoma tumors, and thus can be used as an indicator for
XX      cancer. The invention provides polynucleotides (see AAM88469-70)
XX      encoding Zsig9 polypeptides (see AAM88469-77) including mature
XX      polypeptides, other processed forms, variants and the mouse
XX      orthologues. The Zsig9 gene, or probes derived from it, can be
XX      used to determine if Zsig9 is present on chromosome 10, and if a
XX      mutation has occurred. Antibodies raised against Zsig9 can be
XX      used as diagnostic agents to determine the presence of Zsig9, and
XX      thus the presence of cancer. They can also be labelled with
XX      radioisotopes or fused with toxins and used to treat tumours
XX      which overexpress Zsig9. Antisense nucleotides derived from Zsig9
XX      cDNA can also be used to inhibit the growth of tumour cells. Zsig9
XX      proteins can be used to enhance the growth or development of the
XX      placenta, heart or liver.
XX
XX      Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. NO.:      8.05e-102      Length:      1069
XX      Score:      959.00      Matches:      182
XX      Percent Similarity:      100.00%      Conservative:      0
XX      Best Local Similarity:      100.00%      Mismatches:      0
XX      Query Match:      100.00%      Indels:      0
XX      DB:      20      Gaps:      0
XX
XX      US-10-082-502-19 (1-182) x AAX06970 (1-1069)
XX
XX      QY      1 MetLysGlyTyrGlyTyrPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrrPa 20
XX      |||||||
XX      DB      358 ATGAAAGCGCTGGGGTGGCTAGCCCTACTTTGGGGGTCTGCTGGAACTGGCGT 417
XX      |||||||
XX      QY      21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrr 40
XX      |||||||
XX      DB      418 CGAAGGAGCCAGAGTCACTGAGCTGGAGCTTGGAGGCTCTGGGAGAAATGAGTGG 477
XX      |||||||
XX      QY      41 GluIleAlaArgValAspProLysLeuThrIleGlnMetGlySerPheArgIleAsnPro 60
XX      |||||||
XX      DB      478 GAAATGGCCCGCTGGAGCCCAAGAGACCATTCAGATGGGATCTTCGAAATCAATCCA 537

```

```

QY      61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
XX      |||||||
DB      538 GATGGCAGCAGCTCAGTTGTGGAGGTACTTATGCCCGCTGAGAGCCACCTCAGAG 597
XX      |||||||
QY      81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100
XX      |||||||
DB      598 TTGCTTGAGCAGCGTGTGGACCGAATGAGAGTGGGGAACAGATTGACCTTCTACG 657
XX      |||||||
QY      101 HisArgLysAsnTyrValArgValSerArgAsnGlyLysSerSerGluLeuAspLeu 120
XX      |||||||
DB      658 CACCCCAAGACTAGCTAGCTACGCTCTGAGCCGAGATGGAGAAATCCAGTAAGACTTA 717
XX      |||||||
QY      121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
XX      |||||||
DB      718 CAGGCGATCCGAAATGACTCAGATATCAGCGGACCCCTCAAGTTGCGTGTGAGAGCAT 777
XX      |||||||
QY      141 ValGluGluTyrGluAspGluLeuLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160
XX      |||||||
DB      778 GTGGAAGATACGAGGATGAGCTTATCGAATTCTTCCAGAGAGGCTGACAAACCTTAA 837
XX      |||||||
QY      161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180
XX      |||||||
DB      838 GACAAACTTTGCAGTAAACGCGACAGATCTATGTGACCATGCCCTGCACAGATCTCAGCAT 897
XX      |||||||
QY      181 GluLeu 182
XX      |||||||
DB      898 GAGCTA 903
XX
XX      RESULT 2
XX      ID      AAZ08294
XX      AC      AAZ08294;
XX      AC      AAZ08294;
XX      DT      07-FEB-2000 (first entry)
XX
XX      Mouse ortholog gene encoding Zsig9 secretory protein.
XX
XX      DE
XX      KW      Secretory protein-9; mouse Zsig9; ortholog; overexpression; antagonist;
XX      antibody; antisense nucleotide; tumour; treatment; receptor;
XX      KW      radio-label; polypeptide toxin; down-regulation; diagnostic;
XX      KW      therapeutic; probe; cancer; brain; liver; detection; stomach;
XX      KW      lymphoma; ds.
XX
XX      OS      Mus musculus.
XX
XX      FT      Key
XX      FT      CDS
XX      FT      Location/Qualifiers
XX      FT      sig_peptide
XX      FT      /tag= a
XX      FT      /product= "Mouse Zsig9 secretory protein"
XX      FT      /note= "Overexpressed in tumours"
XX      FT      358..417
XX      FT      mat_peptide
XX      FT      /*tag= b
XX      FT      418..903
XX      FT      /*tag= c
XX      FT      /label= Mature_Zsig9_protein_variant-4
XX
XX      MO9960405-A1.
XX
XX      PD      25-NOV-1999.
XX
XX      XX      19-MAY-1999;      99WO-US11107.
XX      XX      19-MAY-1998;      98US-0081183.
XX
XX      (ZYMO ) ZYMOGENETICS INC.
XX
XX      Moore EE, Taft DW;
XX      WPI; 2000-039447/03.
XX      P-PSDB; AA15136.
XX

```



```

Db 121 GAATTGCCAGCTGGAGCCCAAGAACACATTCAGATGGATCTTTCCGATCAATCCA 180
Oy 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
Db 181 GATGCGACGACGACTCAGTGGTGGAGGTGCTTATGCCCCGCTGACAGGCCACCTCAGAG 240
Oy 81 LeuLeuGluGluValCysAspArgMetGlySGluTyrGlyGluGlnHisLeuProSerThr 100
Db 241 CTGCTGGAGGAGATATGTCAGCCGATGAGAGGTATGGGGAACATGATCTCTCCACC 300
Oy 101 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 120
Db 301 CATCCCAAGAACTACTACGTACTGCTACTGGCCGCAATGCAAGATCCAGTACTGCACTTA 360
Oy 121 GlnGlyIleArgGlyLeuAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 361 CAAGGCATCCGAAATGCACTCAGATATGAGCGCACCCCTCAAGTTTGCGGTGAGAGCAT 420
Oy 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160
Db 421 GTGGAGGAATACGAGATGAGAACTCATTAATCTTTCCCGAGGCTGACAAATGTTAA 480
Oy 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180
Db 481 GACAAACTTTGCAGTAACGAAACAGATCTTTGTGACATGCCCTGCACATATCGCATGAT 540
Oy 181 GluLeu 182
Db 541 GAGCTA 546

```

```

RESULT 4
AA06969
ID AA06969 standard; cDNA: 806 BP.
AA06969;

```

```

XX 10-MAY-1999 (first entry)

```

```

DE Human secretory peptide-9 (Zs19) variant cDNA.

```

```

KW Secretory peptide-9; Zs19; human; tumour marker; cancer; therapy;
diagnosis; growth enhancer; variant; ss.

```

```

XX Homo sapiens.

```

```

FH Key Location/Qualifiers

```

```

FT CDS /*tag= a
104..652

```

```

FT sig_peptide /*tag= b
104..163

```

```

FT mat_peptide /*tag= c
164..649

```

```

FT

```

```

XX MO9901554-A1.

```

```

XX 14-JAN-1999.

```

```

XX 02-JUL-1998; 98WO-US13859.

```

```

XX 17-JUN-1998; 98US-0099005.

```

```

XX 03-JUL-1997; 97US-0051704.

```

```

XX 03-JUL-1997; 97US-0888088.

```

```

XX 19-MAY-1998; 98US-0081338.

```

```

XX 19-MAY-1998; 98US-0085983.

```

```

XX 17-JUN-1998; 98US-0089899.

```

```

XX (ZYMO) ZYMOGENETICS INC.

```

```

XX Jaspers SR, Jelinek LT, Sheppard PO, Whitmore TE;

```

```

PT New mammalian secretory peptide-9 (Zs19) - used as a growth
PT enhancer for placenta, liver and heart, and as an indicator of
PT cancer

```

```

PS Claim 2; Page 73-74; 85pp; English.

```

```

CC This cDNA clone encodes human secretory peptide-9, or Zs19,
CC variant (see AAW88474). Zs19 (see also AAW88469) is overexpressed in
CC human brain, liver, lung, oesophageal, stomach, colon, rectal,
CC thyroid and lymphoma tumors. Thus, Zs19 can be used as an
CC indicator for cancer. Zs19 cDNA was discovered in a placenta
CC clone from a full-term pregnancy cDNA library which contained an
CC expressed sequence tag (see AA06971). The invention provides
CC polynucleotides (see AA06968-70) encoding Zs19 polypeptides (see
CC AA06969-77) including mature polypeptides, other processed forms,
CC variants and mouse orthologues. The Zs19 gene, or probes derived
CC from it, can be used to determine if Zs19 is present on chromosome
CC 12, and if a mutation has occurred. Antibodies raised against
CC Zs19 can be used as diagnostic agents to determine the presence of
CC Zs19, and thus the presence of cancer. They can also be labelled
CC with radioisotopes or fused with toxins and used to treat tumours
CC which overexpress Zs19. Antisense nucleotides derived from Zs19
CC cDNA can also be used to inhibit the growth of tumour cells. Zs19
CC proteins can be used to enhance the growth or development of the
CC placenta, heart or liver.

```

```

XX Sequence 806 BP; 229 A; 183 C; 232 G; 160 T; 2 other;

```

```

Alignment Scores:

```

```

Pred. No.: 1,53e-99 Length: 806
Score: 938.00 Matches: 177
Percent Similarity: 98.35% Conservative: 2
Best Local Similarity: 97.25% Mismatches: 3
Query Match: 97.81% Indels: 0
DB: Gaps: 0

```

```

US-10-082-502-19 (1-182) x AA06969 (1-806)

```

```

Oy 1 MetLysGlyTyrPglYTrPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20
Db 104 ATGAAGAGCTGGGGTGGCTGGCTGCTTCTGGGGGGCCCTGGGAACCGCGTGGCT 163
Oy 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGlu 40
Db 164 CGAGAGACCAAGATCTCCACTGTGAGACATGCAGGCTGTGGTGAATGAATGAAG 223
Oy 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60
Db 224 GAATTTCCCAAGCTGACCCCAAGAACACATTCAGATGGATCTTCTCCGATCAATCCA 283
Oy 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
Db 284 GATGCGACGACGACTCAGTGGTGGAGGTGCTTATGCCCCGCTGAGAGGCCACCTCAGAG 343
Oy 81 LeuLeuGluGluValCysAspArgMetGlySGluTyrGlyGluGlnHisLeuProSerThr 100
Db 344 CTGCTGGAGGAGATATGTCAGCCGATGAGAGGTATGGGGAACACATGATCTCTCCACC 403
Oy 101 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 120
Db 404 CATCCCAAGAACTACTACGTACTGCTACTGGCCGCAATGCAAGATCCAGTACTGCACTTA 463
Oy 121 GlnGlyIleArgGlyLeuAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 464 CAAGGCATCCGAAATGCACTCAGATATGAGCGCACCCCTCAAGTTTGCGGTGAGAGCAT 523
Oy 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160
Db 524 GTGGAGGAATACGAGATGAGAACTCATTAATCTTTCCCGAGGCTGACAAATGTTAA 583
Oy 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180
Db 584 GACAAACTTTGCAGTAACGAAACAGATCTTTGTGACATGCCCTGCACATATCGCATGAT 643

```

Db	Qy	181	Glutru	182	
Db	644	GAGCTA	649		
RESULT 5					
AAZ08293	ID	AAZ08293	standard; DNA; 806 BP.		
XX	AC	AAZ08293;			
XX	DT	07-FEB-2000	(first entry)		
XX	XX	Human Zs19 gene encoding secretory protein variant-4.			
XX	XX	Secretory protein-9; Human Zs19; chromosome 12q15 region; variant;			
XX	XX	overexpression; antagonist; antibody; antisense nucleotide; tumour;			
XX	XX	treatment; receptor; radio-label; fusion; polypeptide toxin; technique;			
XX	XX	down-regulation; probe; diagnostic; therapeutic; cancer; brain; liver;			
XX	XX	detection; stomach; lymphoma; alternative splicing; allelic variation;			
XX	XX	silent mutation; ds.			
XX	OS	Homo sapiens.			
XX	OS	Synthetic.			
XX	FT	Key	Location/Qualifiers		
XX	FT	CDS	104..652		
XX	FT	/*tag= a	/product= "Zs19 secretory protein variant-4"		
XX	FT	/*note= "Overexpressed in tumours"			
XX	FT	104..163	/*tag= b		
XX	FT	mat_peptide	164..649		
XX	FT	/*tag= c	/label= Mature_Zs19_protein_variant-4		
XX	FT	W09360405-A1.			
XX	PD	25-NOV-1999.			
XX	PF	19-MAY-1999;	99WO-US11107.		
XX	PR	19-MAY-1998;	98US-0081183.		
XX	PA	(ZYMO ) ZYMOGENETICS INC.			
XX	PI	Moore EE, Taft DW;			
XX	DR	WPI: 2000-039447/03.			
XX	PS	P-PSDB; AAV15135.			
XX	PT	Detecting tumors using antibodies, antagonists and antisense			
XX	PT	nucleotides to secretory protein-9 (Zs19) -			
XX	PS	Disclosure; Page 33-35; 45pp; English.			
XX	CC	The present DNA sequence is a gene encoding the variant of the secretory			
XX	CC	protein-9, Zs19 that arises due to alternative splicing, allelic			
XX	CC	variation or silent mutations that result in amino acid changes. This			
XX	CC	sequence is mapped to the human chromosome 12q15 region. It is			
XX	CC	overexpressed in tumours. Antagonists, antibodies and antisense			
XX	CC	nucleotides to Zs19 are useful for detecting and treating tumours. The			
XX	CC	antagonist may be an antibody or receptor to Zs19 and it may be radio-			
XX	CC	labelled or fused to a polypeptide toxin. It can be used for down			
XX	CC	regulating the overexpression of Zs19. The gene sequence can be used as			
XX	CC	nucleic acid probes to detect RNA encoding Zs19. The Zs19 sequence			
XX	CC	facilitates improved diagnostic and therapeutic techniques for detecting			
XX	CC	and treating cancers, especially of the brain, liver, stomach, lymphoma,			
XX	CC	etc., at an early stage.			
XX	SO	Sequence 806 BP; 229 A; 183 C; 233 G; 160 T; 1 other;			
XX	XX	Alignment Scores:			

XX 27-APR-1999; 99WO-JP02226.  
 XX 28-APR-1998; 98JP-0119395.  
 XX (SAGA ) SAGAMI CHEM RES CENT.  
 PA (PROT-) PROTEGENE INC.  
 XX Kato S, Kimura T;  
 DR WPI: 2000-023358/02.  
 DR P-PSDB: AAY52391.  
 XX  
 PT Human proteins with transmembrane domains, involved in control of cell  
 PT proliferation and differentiation, useful for treating e.g. cancer or  
 PT inflammation  
 PS Claim 4; Page 106-107; 114pp; English.  
 CC This sequence represents the human cDNA clone HP10390  
 CC which encodes a 20 kD protein with one putative transmembrane  
 CC domain in the N-terminus. The cDNA was isolated from a human stomach  
 CC cancer cell line cDNA library. The protein has no homology with any  
 CC known protein. The protein may be used to raise specific antibodies, as  
 CC assay reagents, as diagnostic tissue markers, for the isolation of  
 CC cognate receptors, ligands and binding proteins, and as biologically  
 CC active agents. Nucleotides encoding the protein may be used as primers  
 CC and probes or antisense molecules, and in gene therapy. Cells transformed  
 CC with these nucleotides may be used to screen for agonists and antagonists  
 CC which are potentially useful therapeutically.

SO Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.55e-99 Length: 814  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3  
 Query Match: 97.81% Indels: 0  
 DB: 21 Gaps: 0

US-10-082-502-19 (1-182) x AA238327 (1-814)

QY 1 MetysglttppgltytrpleuallaleuLeuUgltYValLeuUgltYThAlATrPAla 20  
 DB 145 ATGAAAGGCTGGGTGGCTGGCCCTGCTTCTGGGGCCCTGGGAAACCGCTGGGCT 204  
 QY 21 ArgArSerglnAAspLeuHIsCysGlyAlaCysArGAlaLeuValAspGluLeuGltTrp 40  
 DB 205 CGGAGGACCGAGATCTCCACTGTGGACATCGAGGCTCTGGTGATGAATCGAATGG 264  
 QY 41 GlutleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 265 GAAATTCGCCAGGTGGAGCCCAAGAACATTCAGATGAGATCTTCCGATCAATCCA 324  
 QY 61 AspGlySerGlnSerValValGluValProtyrAlaArgSerGluAlaHisLeuThrIu 80  
 DB 325 GATGGCACCCGCTAGTGGTGGAGTGGCTTATGCCCTCAGAGGCCACCTCACAAGAG 384  
 QY 81 LeuLeuGluGluValAlaCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThr 100  
 DB 385 CTGCTGGAGGAGATATGTGACCGGATGAGAGATATGGGAGAACGATTTGCTTCCACC 444  
 QY 101 HisArgLysAsnTyrrValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 120  
 DB 445 CATGCAAGAACTACGTACGTGTAGTGGCGGAAATGAGAAATCAGTAATCGAGCTA 504  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 505 CAAGGATCCCAATCGACTCAGATATTAGCCGACCCCAAGTTGGCTGTGAAGACATTT 564  
 QY 141 ValGluGluTrpGluAspGluLeuIleGlnPheSerArgGluAlaAspAsnValLys 160

DB 565 GTGAGGAATACGAGGATGAACCTGTAATCTTTCCGAGAGGCTGACAAATGTTAAA 624  
 QY 161 AsplysleuCySserLysArgThrAspLeuCySAspHIsAlaLeuHIsArgSerHisAsp 180  
 DB 625 GACAAACTTTGCAGTAACGAGAACAGATCTTTGTGACCATGCCCTGCACATATCGCATAT 684  
 QY 181 GluLeu 182  
 DB 685 GAGCTA 690  
 RESULT 7  
 ID ABR52765 standard; cDNA; 814 BP.  
 AC ABR52765;  
 XX 15-AUG-2002 (first entry)  
 DE CDNA encoding transmembrane protein 4, a cancer-linked protein.  
 KW Expressed sequence tag; EST; human; cancer; anti-neoplastic;  
 KW cytotoxic T lymphocyte; chemotherapy; cytostatic; chromosome 12;  
 KW transmembrane protein 4; gene; ss.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 145..693  
 FT /\*tag= a  
 FT /product= "Transmembrane protein 4"  
 FN WO200231198-A2.  
 PN 18-APR-2002.  
 PD 11-0CT-2001; 2001WO-US31607.  
 XX 11-0CT-2000; 2000US-239294P.  
 PF 11-0CT-2000; 2000US-239297P.  
 PR 11-0CT-2000; 2000US-239605P.  
 PR 11-0CT-2000; 2000US-239802P.  
 PR 12-0CT-2000; 2000US-239805P.  
 PR 12-0CT-2000; 2000US-239806P.  
 PR 16-0CT-2000; 2000US-240622P.  
 PR 19-0CT-2000; 2000US-241682P.  
 PR 19-0CT-2000; 2000US-241723P.  
 PR 31-0CT-2000; 2000US-244932P.  
 XX (AVAIL-) AVALON PHARM.  
 XX Young PE, Horrigan S, Weaver Z, Endress GA;  
 PI WPI: 2002-463271/49.  
 DR P-PSDB: AAU97063.  
 DR  
 XX  
 PT Identifying modulators of a cancer-related gene to screen agents for  
 PT preventing or treating cancer comprises detecting a difference in the  
 PT expression of cancer-linked genes in the presence or absence of test  
 PT compounds  
 PS Claim 1; Page 43; 66pp; English.  
 XX  
 CC The invention relates to modulators of a cancer-related genes. Also  
 CC described are: (1) processes for identifying an anti-neoplastic agents  
 CC comprising contacting a cell exhibiting neoplastic activity with a  
 CC compound first identified as a cancer related gene modulator, and  
 CC detecting a decrease in the neoplastic activity. (2) a process for  
 CC determining the cancerous state of a cell by determining an increase in  
 CC the level of expression of at least one gene, where an elevated  
 CC expression relative to a known non-cancerous cell indicates a cancerous  
 CC state or potentially cancerous state. The anti-neoplastic agent is  
 CC useful for treating cancer or for protecting an animal against cancer.  
 CC The immunogenic composition is also useful for treating cancer in an



CC animal, where the composition elicits the production of cytotoxic T  
 CC lymphocytes specific for the immunogenic composition. Preferably, the  
 CC animal is a human. The cancer-linked genes and polypeptides are also  
 CC useful as targets for cancer therapy or chemotherapy. The present  
 CC sequence represents a cancer-linked gene located on chromosome 12,  
 CC which encodes transmembrane protein 4.

Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Alignment Scores:

Pred. No.: 1.55e-99 Length: 814  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3  
 Query Match: 97.81% Indels: 0  
 DB: 24 Gaps: 0

US-10-082-502-19 (1-182) x ABK52765 (1-814)

QY 1 MetlysglytrpGLYtrpLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
 DB 145 ATGAAGGCTGGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 204  
 QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 DB 205 CGAGAGAGCCAGAGATCTCCAGTGTGAGCATGAGAGGCTCTGGTGGATGAATGATG 264  
 QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 265 GAAATGGCCAGGTGGACCCAGAGAACATTCAGATGGATGCTTCCGATCAATCCA 324  
 QY 61 AspGlySerGlnSerValAlaGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 80  
 DB 325 GATGGACAGCCAGTGTGAGTGGAGTGGCTTATGCCCCCTCAGAGGCCACCTTACAGAG 384  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThr 100  
 DB 385 CTGCTGGAGAGATATGTGACCGGATGAGAGATGATGGAGAACAGATTGATCTCCACC 444  
 QY 101 HisArgLysAsnTrpValArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
 DB 445 CATGGCAAGAACTACGTCGTCGTGTGAGGCGCGAATGAGAACTGAACTGGACCTA 504  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 505 CAAGGCATCCGAATCGACTGAGATATTAGGGCCACCTCAAGTTGGCTGTGAGAGCAT 564  
 QY 141 ValGluGluTrpGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
 DB 565 GTGGAGGAATACGAGATGAACTGAACTTCTTCCCGAGAGGCTGACAAATGTTAAA 624  
 QY 161 AspLysLeuCysSerLysArgTrpAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 DB 625 GACAAACTTTGCGATGAGCAAGACAGATCTTTGTGACCATGCCCTGACATATGCGCAGAT 684  
 QY 181 GluLeu 182  
 DB 685 GAGCTA 690

RESULT 8  
 ABL41995  
 ID ABL41995 standard; DNM: 814 BP.

AC ABL41995;

DT 11-JUN-2002 (first entry)

DE Nucleotide sequence of human polypeptide HP10390.

KW Antibody; antigen; transmembrane domain protein; HP10390; gene; ss.

OS Homo sapiens.

Key Location/Qualifiers  
 CDS 145..693  
 /tag= a  
 /product= "HP10390"

MO200208416-A1.

31-JAN-2002.

24-JUL-2001; 2001WO-JP06371.

24-JUL-2000; 2000JP-0222743.

24-AUG-2000; 2000JP-0254407.

(NISC-) JAPAN SCI & TECHNOLOGY CORP.

Kato S, Nagata N, Fujimura N, Kobayashi M, Ito K, Ishizuka Y;

WPI; 2002-195877/25.

P-PSDB; ABB09717.

Antibody preparation by inoculation of an animal with a vector  
 expressing a fusion protein of an antigen on the C-terminal side of a  
 transmembrane domain for use as drugs, diagnostic reagents and  
 laboratory reagents

Example; Page 33-35; 45pp; Japanese.

The specification describes a method of antibody preparation. The  
 method comprises inoculating an animal with a vector expressing  
 a fusion protein having an antigen protein fused to the C-terminal  
 side (extracellular) of a transmembrane domain protein (the  
 C N-terminal side of which is intracellular), and then isolating and  
 purifying the antibody from the animal. The antibodies can be used  
 as drugs, vaccines, diagnostic reagents and laboratory reagents. The  
 present sequence encodes a polypeptide, designated HP10390, which was  
 used in the course of the invention.

Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Alignment Scores:

Pred. No.: 1.55e-99 Length: 814  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3  
 Query Match: 97.81% Indels: 0  
 DB: 24 Gaps: 0

US-10-082-502-19 (1-182) x ABL41995 (1-814)

QY 1 MetlysglytrpGLYtrpLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
 DB 145 ATGAAGGCTGGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 204  
 QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 DB 205 CGAGAGAGCCAGAGATCTCCAGTGTGAGCATGAGAGGCTCTGGTGGATGAATGATG 264  
 QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 265 GAAATGGCCAGGTGGACCCAGAGAACATTCAGATGGATGCTTCCGATCAATCCA 324  
 QY 61 AspGlySerGlnSerValAlaGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 80  
 DB 325 GATGGACAGCCAGTGTGAGTGGAGTGGCTTATGCCCCCTCAGAGGCCACCTTACAGAG 384  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThr 100  
 DB 385 CTGCTGGAGAGATATGTGACCGGATGAGAGATGAGTGGGAAACAGATTGATCTTCCACC 444  
 QY 101 HisArgLysAsnTrpValArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
 DB 445 CATGGCAAGAACTACGTCGTCGTGTGAGGCGCGAATGAGAACTGAACTGGACCTA 504

QY 121 GlnGlyIleAArgIleAspSerAspIleSerGlyThrLeuAlaCysGluSerIle 140  
DB 505 CAAAGCATCCGATCGACTGATATAGCCGACCCCTCAAGTTGCGTGTAGAGCAT 564  
QY 141 ValGluGluTyrGlnAspGluLeuIleGluPhePheSerArgGluAlaAspAsnVal 160  
DB 555 GTGAGAGAAATACGAGATGAACTCATTAATCTTTCCCGAGAGCTGACAAATGTTAA 624  
QY 161 AspIleAspCysSerIleArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
DB 625 GACAAACCTTTCAGTAAACGAAACAGATCTTTGTGACATGCCCTGCACATATCCATGAT 684  
QY 181 GluLeu 182  
DB 685 GAGCTA 690  
RESULT 9  
ID ABR09772 standard; cDNA; 814 BP.  
XX ABR09772:  
AC 14-MAR-2002 (first entry)  
DE Human ovarian tumour protein encoding cDNA #305.  
XX Human ovarian tumour protein; cancer; cytostatic; immunostimulant; ss;  
KW gene therapy; CD4+ T cell; CD8+ T cell; PCR primer.  
OS Homo sapiens.  
XX MO200190154-A2.  
PN 29-NOV-2001.  
PD 23-MAY-2001; 2001WO-US16895.  
PE 24-MAY-2000; 2000US-207107P.  
PR 13-JUN-2000; 2000US-211457P.  
PR 21-JUN-2000; 2000US-213673P.  
PR 03-AUG-2000; 2000US-223288P.  
PR 01-MAR-2001; 2001US-272790P.  
XX (CORI-) CORIAX CORP.  
XX Xu J, Mitcham JL, Harlocker SL, Dillon DC, Secrist H, Lodes MJ;  
PI Algate PA, Fling SP, Mannion J, Benson DR, Carter D;  
DR WPI: 2002-097641/13.  
XX New isolated polynucleotide encoding polypeptide comprising portion of  
PT ovarian tumour protein, useful for detection, diagnosis and therapy of  
PT human ovarian cancer -  
PS Claim 1; Page 253-254; 285bp; English.  
XX The invention relates to an isolated polynucleotide encoding a  
CC polypeptide comprising a portion of an ovarian tumour protein. The  
CC sequences of the invention are useful for stimulating an immune response  
CC and for treating ovarian cancer in a patient. An antigen presenting cell  
CC that expresses the sequences is useful for treating ovarian cancer by  
CC incubating CD4+ and/or CD8+ T cells isolated from a patient. The T cells  
CC can then be proliferated and administered to the patient to inhibit the  
CC development of cancer. The DNA sequences are useful as probes or primers  
CC for nucleic acid hybridisation, to direct expression of a polypeptide in  
CC appropriate host cells. Detecting the presence of a cancer in a patient  
CC involves obtaining a biological sample from the patient, contacting the  
CC biological sample with an agent that binds to the protein, detecting the  
CC amount of protein that binds to the agent, comparing the amount of  
CC protein to a predetermined cut-off value and determining the presence of  
CC cancer. Sequences ABR09464-ABR09802 represent PCR primers and cDNA  
CC molecules encoding ovarian tumour proteins of the invention.

XX SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;  
Alignment Scores:  
Pred. No.: 1.55e-99 Length: 814  
Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: Gaps: 0  
US-10-082-502-19 (1-182) x ABR09772 (1-814)  
QY 1 MetIysGlyTyrPglYThrLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
DB 145 ATGAAAGCTGGGGTGTGGCTGGCCCTGCTTCTGGGGGCGCTCTGGAAACCGCTGGGCT 204  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
DB 205 CGAGAGACCCAGATCTCCACTGTGGACATGCACGCCCTCTGTGGATGAACTAGAAATGG 264  
QY 41 GluIleAlaArgValAspProValYsrThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 265 GAAATTCGCCAGGTGAGACCCCAAGAACCATTCAGATGGGATCTTCCGGATCAATCCA 324  
QY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 325 GATGCGACCGCATGCTAGTGTGGAGTGCCTATAGCCCGCTGAGAGCCACCTCAGAG 384  
QY 81 LeuLeuGluGluValCysAspArgMetIysGluTyrGlyGluGlnIleAspProSerThr 100  
DB 385 CTGCTGGAGAGATATGTTGGACCGGATGAGAGATGTGGGAAACATGATGATCTTCCACC 444  
QY 101 HisArgIysAsnTyrValArgValValSerArgAsnGlyGluSerSerGluLeuAspLeu 120  
DB 445 CATGCCAAGAACTAGCTACGTACGTAGTGGCCCGGATGAGAAATCCAGTAAGTGAACCTA 504  
QY 121 GlnGlyIleAArgIleAspSerAspIleSerGlyThrLeuAlaCysGluSerIle 140  
DB 505 CAAAGCATCCGATCGACTGATATAGCCGACCCCTCAAGTTGCGTGTAGAGCAT 564  
QY 141 ValGluGluTyrGlnAspGluLeuIleGluPhePheSerArgGluAlaAspAsnVal 160  
DB 565 GTGAGAGAAATACGAGATGAACTCATTAATCTTTCCCGAGAGCTGACAAATGTTAA 624  
QY 161 AspIleAspCysSerIleArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
DB 625 GACAAACCTTTCAGTAAACGAAACAGATCTTTGTGACATGCCCTGCACATATCCATGAT 684  
QY 181 GluLeu 182  
DB 685 GAGCTA 690  
RESULT 10  
ID AAX22112 standard; DNA; 1180 BP.  
XX AAX22112:  
AC 18-MAY-1999 (first entry)  
DE Human secreted protein gene 2 clone H2MBB56.  
XX Human; secreted protein; gene therapy; protein therapy; tissue; cancer;  
KW tumour; neurodegenerative disorder; leukaemia; autoimmune disease; AIDS;  
KW developmental abnormality; foetal deficiency; Alzheimer's disease;  
KW cognitive disorder; schizophrenia; immunological disorder; mood disorder;  
KW immune deficiency disease; respiratory disorder; arthritis; skeletal;  
KW haematopoietic disorder; neural; osteoporosis; metabolic disorders;  
KW cardiovascular; endocrine; gastrointestinal; asthma; diagnosis; ds.  
XX Homo sapiens.

PN MO9901020-A2.  
 XX 14-JAN-1999.  
 PD  
 XX 30-JUN-1998; 98WO-US13608.  
 XX  
 PR 12-SEP-1997; 97US-0058663.  
 PR 01-JUL-1997; 97US-0051381.  
 PR 01-JUL-1997; 97US-0051480.  
 PR 12-SEP-1997; 97US-0058598.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Carter KC, Endress GA, Feng P, Rosen CA, Ruben SM;  
 DR WPI: 1999-105683/09.  
 DR P-PSDB; AAY01136, AAY01162, AAY01163.  
 XX  
 XX  
 PT New isolated human genes and the secreted polypeptides they encode -  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, immune deficiency diseases or blood  
 XX disorders  
 PS Claim 4; Page 117; 179pp; English.  
 XX  
 CC The invention relates to nucleic acid sequences (AA22111 to AA22134)  
 CC encoding human secreted proteins (AAY01135 to AAY01158). The secreted  
 CC protein gene sequences are deposited with the ATCC under deposit number  
 CC ATCC 209118. Host cells comprising recombinant vectors containing the  
 CC nucleic acid sequences are used for the recombinant production of the  
 CC secreted proteins. The polynucleotide and amino acid sequences of the  
 CC for are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Pathological conditions can  
 CC be also diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new  
 CC polynucleotides. Specific uses are described for each of the  
 CC polynucleotides, based on which tissues they are most highly expressed  
 CC in, and include developing products for the diagnosis or treatment of  
 CC cancer, tumours, developmental abnormalities and cognitive disorders,  
 CC autoimmune diseases, lymphomas, Alzheimer's and cognitive disorders,  
 CC schizophrenia, immunological disorders, immune deficiency diseases  
 CC (AIDS), mood disorders, respiratory disorders, arthritis, asthma,  
 CC hematopoietic disorders, neural disorders, skeletal disorders,  
 CC osteoporosis, metabolic disorders, cardiovascular disorders, endocrine  
 CC disorders or gastrointestinal disorders. The polypeptides are also useful  
 CC for identifying their binding partners. The present sequence represents a  
 CC gene encoding a human secreted protein (see descriptor line for gene  
 CC number and clone identification).  
 CC  
 CC  
 SQ Sequence 1180 BP; 258 A; 304 C; 363 G; 250 T; 5 other;  
 Alignment Scores:  
 Pred. No: 2.54e-99 Length: 1180  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3  
 Query Match: 97.81% Indels: 0  
 DB: 20 Gaps: 0  
 US-10-082-502-19 (1-182) x AA22112 (1-1180)  
 QY 1 MetysgLyTrpGlyTrpLeuAlaLeuLeuLeuValLeuLeuGlyThrAlaTrpAla 20  
 DB 513 ATGAAGGCGGGGTGGCGCCCTCTCTCTGCGGCGCTGCGGCAACCGCTGGCT 572  
 QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 DB 573 CGSAGAGAGCCAGATCTCCACTGTGAGCATGCGAGCGCTCTGTGATGAACAGATG 632  
 QY 41 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 633 GAAATTCGCCAGGTGACCCCAAGAACATTCAATGATGATGATGATGATGATGATGAT 692

QY 61 AspGlySerGlnSerValValGluValProTyrAlaIleArgSerGluAlaHisLeuThrGlu 80  
 DB 693 GATGGAGCCACTGATGAGTGGAGGTGGCTTATGCCCTCAGAGCCACCTCAGAG 752  
 QY 81 LeuLeuGluGluValAlaCysAspArgMetLysGlyTyrGlyGlnIleAspProSerThr 100  
 DB 753 CTGCTGAGAGGATATGACCGGATGAAAGATATGGGAAACGATTGATCTTCCACC 812  
 QY 101 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 120  
 DB 813 CATGCCAAGAACTACGTACGTGTAGTGGCCGGAATGAGATTCAGTGAACCTGACCTA 872  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 873 CAAGCATCCCAATCGACTGAGATTTAGCGGCACCTCAAGTTTGGCTGAGACGATT 932  
 QY 141 ValGluGluTyrGlyAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
 DB 933 GTGGAGAAATACGAGATGAACTCATTTGATTTCCCGAGAGGCTGACAAATGTTAA 992  
 QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 DB 993 GACAAACTTTCAGTAAGCGAAACAGATCTTTGTGACCATGCCCCCTGCACATATGCATGAT 1052  
 QY 181 GluLeu 182  
 DB 1053 GAGCTA 1058  
 RESULT 11  
 AAS21330  
 ID AAS21330 standard; cDNA; 1210 BP.  
 XX  
 AC AAS21330;  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human cDNA sequence encoding for PRO4426 polypeptide.  
 XX  
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
 KW breast; prostate; cervical; tumour necrosis factor- $\alpha$ ; TNF- $\alpha$ ;  
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
 KW adipocyte; A-peptide; factor VIIa; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200140466-A2.  
 XX  
 PD 07-JUN-2001.  
 XX  
 PF 01-DEC-2000; 2000WO-US32678.  
 XX  
 PR 01-DEC-1999; 99WO-US28301.  
 PR 01-DEC-1999; 99WO-US28634.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 02-DEC-1999; 99WO-US28564.  
 PR 02-DEC-1999; 99WO-US28565.  
 PR 09-DEC-1999; 99US-0170262.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 20-DEC-1999; 99WO-US30911.  
 PR 20-DEC-1999; 99WO-US30999.  
 PR 30-DEC-1999; 99WO-US31243.  
 PR 06-JAN-2000; 2000WO-US00277.  
 PR 06-JAN-2000; 2000WO-US00376.  
 PR 11-FEB-2000; 2000WO-US00376.  
 PR 18-FEB-2000; 2000WO-US03365.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 18-FEB-2000; 2000WO-US04342.  
 PR 22-FEB-2000; 2000WO-US04414.  
 PR 24-FEB-2000; 2000WO-US04914.  
 PR 24-FEB-2000; 2000WO-US05004.  
 PR 01-MAR-2000; 2000WO-US05601.  
 PR 21-MAR-2000; 2000WO-US07377.  
 PR 30-MAR-2000; 2000WO-US07532.  
 PR 30-MAR-2000; 2000WO-US08439.



Score: 931.00 Matches: 176  
 Percent Similarity: 97.808 Conservative: 2  
 Best Local Similarity: 96.708 Mismatches: 4  
 Query Match: 97.08% Indels: 0  
 DB: 20 Gaps: 0

US-10-082-502-19 (1-182) x AAK97884 (1-714)

QY 1 MetLysGlyLTPGlyLTPLeuAlaLeuLeuGlyValLeuLeuGlyThraLTPAla 20  
 DB 33 ATGMAAGCTGGGGTGGCTGGCTGGCTGGGGCCCTCTGGGAACCCCTGGGCT 92  
 QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrp 40  
 DB 93 CGAGAGACCCAGGATCTCCACTGTGTGAGCATGAGGCTGTGTGTGATGACTAGAAATGG 152  
 QY 41 GlnIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 153 GAAATTTGCCAGCTGGACCCCAAGAACATTCAGATGGGATCTTCCGATCAATCCA 212  
 QY 61 AspGlySerGlnSerValAlaGluValProTyrAlaAspSerGluAlaHisLeuThrglu 80  
 DB 213 GATGCAAGCACTCACTGTGTGAGGCTGTGTATGCCCTCAGAGGCCCACTCAGACAG 272  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAsnProSerThr 100  
 DB 273 CTGCTGGAGAGATATGTGACCGGATGAGAGATGAGGGAACAGATTGATCTTCCACC 332  
 QY 101 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyLysSerGluLeuAspLeu 120  
 DB 333 CATGCCAAGAACTACGTACGTGTGTGTGGCCCGAATGAGAAATCAGACTGACCTGA 392  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 393 CAAGGCATCCGAATCGACTCGATATTAGCGGCACTCAAGTTGGCTGGGAGCATTT 452  
 QY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
 DB 453 GTGAGAGAAATGAGAGATGAACTCATTTGTTTCCGAGAGGCTGCACATGTTTAA 512  
 QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspIleAlaLeuHisAspSerHisAsp 180  
 DB 513 GACAAACTTTTGCATTAAGCAACAGATCTTTGTGACCATGCTCCGACATATCGCATGAT 572  
 QY 181 GluLeu 182  
 DB 573 GAGCTA 578

RESULT 13  
 ID ABR36007 standard; cDNA: 657 BP.  
 AC ABR36007;  
 DT 08-MAY-2002 (first entry)  
 DE cDNA sequence #398 encoding novel human secreted protein.  
 XX  
 KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;  
 KW Immune deficiency disorder; blood disorder; inflammatory disorder;  
 KW Infectious disorder; allergic condition; neurodegenerative disorder;  
 KW Liver fibrosis; coagulation disorder; gene therapy; antitubercular;  
 KW tumour; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200177289-A2.  
 XX  
 PD 18-OCT-2001.  
 XX  
 PF 29-MAR-2001; 2001WO-US10232.  
 XX  
 PR 06-APR-2000; 2000US-195605P.

XX (GENY ) GENETICS INST INC.  
 PA Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Agostino MJ, Bowman MR, Spaulding V, Wong GG;  
 PT Clark HF, Fechel K, Howes SH, Resnick RJ, Gultokta K, Graham JR;  
 XX WPI: 2002-179322/23.

Six hundred and twenty three polynucleotides derived from a variety of human tissue sources which encode secreted proteins, useful for treating immune deficiencies and disorders such as autoimmune disorders

Claim 1; Page 297; 393pp; English.

CC The present invention relates to the isolation of novel cDNA sequences  
 CC which encode human secreted proteins. The cDNA sequences have been  
 CC derived from a variety of human tissues. The invention also provides  
 CC a method for producing proteins from these polynucleotide sequences.  
 CC The proteins are useful for identifying compounds that modulate their  
 CC activity and production. The sequences of the invention are  
 CC useful for treating diseases such as hyperproliferative disorders  
 CC (e.g. cancer), immune deficiency disorders (e.g. severe combined  
 CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple  
 CC sclerosis), blood disorders (e.g. thrombocytopenia), inflammatory  
 CC disorders (e.g. arthritis), infectious disorders (e.g. hepatitis),  
 CC allergic conditions (e.g. asthma), neurodegenerative disorders (e.g.  
 CC Alzheimer's disease), liver fibrosis, coagulation disorders (e.g.  
 CC haemophilia), and tumours. The polynucleotide sequences of the  
 CC invention are also useful in gene therapy. ABR35610-ABR35623 represent  
 CC the cDNA sequences of the invention that encode for novel human  
 CC secreted proteins.

Sequence 657 BP; 177 A; 149 C; 185 G; 146 T; 0 other;

Alignment Scores:  
 Pred. No.: 8,916-95 Length: 657  
 Score: 896.00 Matches: 170  
 Percent Similarity: 97.19% Conservative: 3  
 Best Local Similarity: 95.51% Mismatches: 5  
 Query Match: 93.43% Indels: 0  
 DB: 24 Gaps: 0

US-10-082-502-19 (1-182) x ABR36007 (1-657)

QY 5 GlyTTPLeuAlaLeuLeuGlyValLeuLeuGlyThraLTPAlaArgArgSerGln 24  
 DB 1 GGTGGCTGGCCCTCTCTGTGGGGCCCTCTGGGAACCCGCTGGGCTCGAGAGCCAG 60  
 QY 25 AspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrpGluIleAlaArg 44  
 DB 61 GATCTCCACTGTGAGCATGAGGCTGTGTGATGAACTAGAAATGGAAATGGCCAC 120  
 QY 45 ValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnProAspGlySerGln 64  
 DB 121 GTGGACCCCAAGAACCATCATGATGGGATCTTCCGAGATCAATCCAGATTGAGCCAG 180  
 QY 65 SerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGluLeuGluGlu 84  
 DB 181 TCAGTGGTGGAGTGGCTTATGCCGCTCAGAGGCCCAACATCAGAGCTGTGGAGAG 240  
 QY 85 ValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsn 104  
 DB 241 ATATGTGACCGGATGAAGAGATATGGGAAACAGATATCTTCCACCATCCAGAAC 300  
 QY 105 TyrValArgValAlaSerArgAsnGlyLysSerSerLysLeuAspLeuGlnGlyIleArg 124  
 DB 301 TACGTACGTGTAGTGGCCGGAATGAGAAATCCAGTACACTGAGACCTACAAAGCATCCGA 360  
 QY 125 IleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGluTyr 144  
 DB 361 ATCGACTCAGATATTACGGCACCCTCAAGTTTCCGTGTGAGAGCAATTTGGAGGAATAC 420

OY 145 GlnAspGluLeuIleGluPheSerArgGluAlaAspAsnValLysAspLysLeuGys 164  
|||||  
Db 421 GAGGATGACACATTCGATTCCTTTCCGACAGCGCTGACAAATGTTAAAGACAAACTTGGC 480  
OY 165 SerLysAspGlnThrAspLeuGysAspHisAlaLeuHisArgSerHisAspGluLeu 182  
|||||  
Db 481 AGTAAAGCAACGATCTTTGTGACCATCCCTGCACATATCGCATGATGAGCTA 534  
RESULT 14  
ID AAX97837 standard; cDNA: 832 BP.  
XX AAX97837;  
AC AAX97837;  
XX 23-SEP-1999 (first entry)  
XX Human secreted protein encoding cDNA #25.  
DE XX  
KM Secreted protein: human; cytosolic; thrombotic; osteopathic; forensic;  
KW diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
XX  
OS Homo sapiens.  
XX MO9925825-A2.  
XX PD 27-MAY-1999.  
XX PF 13-NOV-1998; 98WO-1B01862.  
XX PR 04-SEP-1998; 98US-0099273.  
PR 13-NOV-1997; 97US-0066677.  
PR 17-DEC-1997; 97US-0069957.  
PR 09-FEB-1998; 98US-0074121.  
PR 13-APR-1998; 98US-0081563.  
PR 10-AUG-1998; 98US-0096116.  
XX  
XX (GIST ) GENSET.  
XX PA  
XX PI Bougueleret L, Duclert A, Dumas Milne Edwards J;  
XX DR WPI: 1999-347472/29.  
XX P-PSDB: AAY36153.  
XX PT Extended cDNAs encoding secreted proteins  
XX Example 28; Page 194-195; 307pp; English.  
XX PS  
XX CC AAX97813-X97906 represent extended cDNA's which encode novel human  
CC secreted proteins (see AAY36129-X36222) and which have cytosolic,  
CC thrombotic and osteopathic activity. The extended cDNAs can be used to  
CC express secreted proteins or parts of them or to obtain antibodies  
CC capable of binding to the secreted proteins. They may also be used in  
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC CC also include design of expression vectors and secretion vectors.  
XX  
SO Sequence 832 BP; 224 A; 194 C; 243 G; 168 T; 3 other:  
Alignment Scores:  
Pred. No.: 7.95e-94 Length: 832  
Score: 889.00 Matches: 175  
Percent Similarity: 96.22% Conservative: 3  
Best Local Similarity: 94.59% Mismatches: 4  
Query Match: 92.70% Indels: 3  
DB: 20 Gaps: 0  
US-10-082-502-19 (1-182) x AAX97837 (1-832)  
OY 1 MetLysGluYTrpGluYTrpLeuAlaLeuLeuGluYValLeuLeuGluYThrAlaTrpAla 20  
|||||  
Db 148 ATGAAGAGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 207  
OY 21 ArgArgSerGlnAspLeuHisCysGluAlaCysArgAlaLeuValAspGlu-LeuGluTr 40  
|||||

Db 208 CGGAGAGCCAGGATCTCCACTGTGGACATGACAGGCTGTGGTGAAGAACTAGAAATG 267  
OY 40 pGluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAspPr 60  
|||||  
Db 268 GGAATTCGCCAGGAGGAGCCCAAGACACATTCAGATGGATCTTCCGATCAATCC 327  
OY 60 oAspLysSerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrG1 80  
|||||  
Db 328 AGATGCGAGCCAGTCAGTGGTGGAGCTCTTATGCCGCTCAGAGGCCACCTCACGA 387  
OY 80 uLeuLeuGluGluValCysAspArgMetLysGluYTrpGluGlnIleAspProSerTh 100  
|||||  
Db 388 GCTGCTGGAGGATATGTGACCGGATGAGAGATGAGGAAACGATGATCTCTCCAC 447  
OY 100 rHisArgLysAsnTyrValArgValValSerArgAsnGlyGluSerSerGluLeuAspLe 120  
|||||  
Db 448 CCATCGCAAGAACTACGATCGTGTAGTGGCGGCAATGCAATTCAGTGAATGAGACT 507  
OY 120 uGlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLys-PheAlaCysGluSerI 140  
|||||  
Db 508 ACAAGGCATCCGAATCGACTCAGATATAGCGGCACCTCAAGBTTCGTGGAGCA 567  
OY 140 IeValGluGluTrpGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValL 160  
|||||  
Db 568 TTGTGGAGAAATACGAGATGACATGAAATCTTTCCGAGAGCTGACAAATGTA 627  
OY 160 yAspLysLeuCysSerLysArgThrAspLeuGysAspHisAlaLeuHisArgSer-His 179  
|||||  
Db 628 AAGCAAACTTGGCAGTAAAGCAACAGATCTTTGGACCAATGCCCTGCACATTCGGCAT 687  
OY 180 AspGluLeu 182  
|||||  
Db 688 GATGAGCTA 696  
RESULT 15  
ID AAA43525 standard; cDNA: 592 BP.  
XX AAA43525;  
XX AC  
XX AA43525;  
XX DT 21-AUG-2000 (first entry)  
XX DE  
XX XX Mouse secreted expressed sequence tag SEQ ID NO:100.  
XX  
KW Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;  
KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
KW immunomodulatory; haematopoietic; chemokine; analgesic; haemostatic;  
KW thrombolytic; antiinflammatory; cytosolic; antibacterial; antifungal;  
KW antiviral; antidiabetic; antiasthmatic; vulnery; antiparkinsonian;  
KW antitumor; osteopathic; neuroprotective; noctropic; antiparasitic;  
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
KW central nervous system disorder; Alzheimer's disease; stroke;  
KW Parkinson's disease; Huntington's disease; coagulation disorder;  
KW hemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
KW tumour; infection; depression; psoriasis; ss.  
XX  
XX Mus musculus.  
XX WO200021991-A1.  
XX PD 20-APR-2000.  
XX PF 15-OCT-1999; 99WO-US24206.  
XX PR 15-OCT-1998; 98US-0104436.  
XX PA (GENY ) GENETICS INST INC.  
XX PI Jacobs K, McCoy JM, Lavallee ER, Collins-Racie LA, Evans C;

FI Merberg D, Treacy M, Bowman MR;  
 DR WPI: 2000-317938/27.  
 XX  
 XX Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (SESTs), useful for treating various disorders  
 PT such as autoimmune, infectious, and central nervous system disorders -  
 XX  
 PS Claim 1, Page 228--229; 803pp; English.  
 XX

CC sequence tags (SESTRs). Isolated from human, mouse, chicken and rat  
CC tissue sources. The SESTRs can have a range of activities depending on  
CC the tissues they were isolated from. The activities include:  
CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
CC chemokinetic; angiogenic; haemostatic; thrombolytic; antiinflammatory;  
CC cytostatic; antibacterial; antifungal; antiviral; antibiogenic;  
CC antiastatic; vulnerrary; antitumor; osteopathic; neuroprotective;  
CC neurotropic; antiparkinsonian; antipsoriatic; cerebroprotective;  
CC anticonvulsant; and antidepressant. The SESTRs can be used for gene  
CC therapy and in vaccines. The SESTRs are useful as probes for the  
CC identification and isolation of full-length cDNAs and genomic DNA  
CC molecules which correspond to the SESTRs. Proteins encoded by the SESTRs  
CC are useful in assays for determining biological activity and raising  
CC antibodies. They may be useful for treatment of autoimmune disorders  
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
CC osteoporosis, osteoarthritis, central nervous system disorders  
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
CC disease), tumours, bacterial, fungal or viral infections, depression and  
CC psoriasis. A445526 to A445591 represent linker variants which are given  
CC in the exemplification of the present invention.

sequence 592 BP; 159 A; 139 C; 167 G; 123 T; 4 other,

**Alignment Scores:**

Pred. NO.:	3, 04e-90	592
Score:	86.50	174
Percent Similarity:	95.11%	Conservative: 1
Best Local Similarity:	94.57%	Mismatches: 7
Query Match:	89.31%	Indels: 4
DB:	21	Gaps: 0

OS-10-062-502-19 (1-182) x AAA43525 (1-592)

OY	1	MeLTVSGIYTRGILTYRLeuLeuLA-LeuLeuLeuGIYValLeuLeuGIYThrLAATPAl	20
Db	44	ATGAAGATGGGGGTGGCTGGCTACCCCTACTTTTGGGGGGTCTGCTGGGAACTGGCTGGGA	103
OY	20	AATGATserGIAsnAspLeuHIScySGIYAlAcysArGAlaLeuValAspGIuLeuGIuTr	40
Db	104	TGGAAGGAGCCAAATATCTACACTGTGGANCTTGGCAGGGCTTGGTGTGATGAATTAAGATG	163
OY	40	pGIuLeuLeuArGValAspProGlybYbThrIleGIuMetcYserPheArGIleasnPr	60
Db	164	GGAAATTTCCCGCGGGAGCCCAAGAAAGACCATTCACATGGAGATCTTCGGAATCAATCC	223
OY	60	oAspGIYserGIAsnserValIaGIuValIProTYAlArGserGIuLAsnIshLeuThrGI	80
Db	224	AGATGGCAGCCAGTCAGTGTGGAGAGTACCTTATGCCCGCTCGAAGGCCACCACTCACGA	283
OY	80	uLeuLeuGIuGIuValcYsasparGmetcYbGIuTYrGIYGIuGIuInIleAspProserTh	100
Db	284	GTTCCTGTGAGGAGGTGTGTGACCCAAATGAAGAGATACGGGA-CAGATTGA-CCTTCTAC	341
OY	100	rHisArGIysAsnTYrValArGValIValserArGAsnGIcYserSerGIuLeuAspLe	120
Db	342	CCACCGGACGAACACTACTACCGCTGCTGACCCGGAAATGGAGAAATCCAGTGAACATACTT	401
OY	120	uGIuGIYIleArGIleAspserAspIleSerGIYThrLeuYsPheAlAcysGIuSerI	140
Db	402	ACAGGGCATCCGAATGTACTCAGATTATCANGCGACACCTCTAAATTTGGCTGTGAGAGCAT	461

Qy	140	evalalngluyrygrlumaapglulelllecluphpheseraarglualaapasnvally	167
Db	462	tetggaaamaaatacggagatgagcttatgcattatcttccacagagagcctgacaaacgttaa	521
Qy	160	sasprylsleucysserlysarargthnaspleucysasphialaleuhlsarsser-h1a	180
Db	522	agacaaactttggacataagcggacagatctatgtgaccatccctccacacagatcctcagc	581
Qy	180	spgluleu 182	
Db	582	aaagacctc 589	

Search completed: March 30, 2003, 22:42:55  
Job time : 144.849 secs





GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:24:04 ; Search time 1725.3 Seconds

(without alignments)  
3070.029 Million cell updates/sec

Title: US-10-082-502-19

Perfect score: 959  
Sequence: 1 MKGGMALILGLVLTGTA.....LCSKRDPLCHALHRSDEL 182

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xip  
-O=/cg2\_1/USPRO.spool/US10082502/runat.24032003.135059.6547/app-query.fasta.1.1308  
-DB=Genmb1 -GFW=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=pcr -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USBR=US10082502.ecgn.1.1.4758.grnat.24032003.135059.6547 -NCP=6 -TCP=3  
-NO\_XLPHY -NO\_MAMP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMECUT=120  
-WARN\_TIMECUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :  
Genmb1:.\*  
1: gb\_ba:.\*  
2: gb\_hlg:.\*  
3: gb\_in:.\*  
4: gb\_com:.\*  
5: gb\_ov:.\*  
6: gb\_pat:.\*  
7: gb\_ph:.\*  
8: gb\_pl:.\*  
9: gb\_pr:.\*  
10: gb\_ro:.\*  
11: gb\_sts:.\*  
12: gb\_sy:.\*  
13: gb\_un:.\*  
14: gb\_vl:.\*  
15: em\_ba:.\*  
16: em\_fun:.\*  
17: em\_hum:.\*  
18: em\_in:.\*  
19: em\_mu:.\*  
20: em\_om:.\*  
21: em\_of:.\*  
22: em\_ov:.\*  
23: em\_pat:.\*  
24: em\_ph:.\*  
25: em\_pl:.\*  
26: em\_to:.\*  
27: em\_sts:.\*  
28: em\_un:.\*

29: em\_vl:.\*  
30: em\_hlg\_hum:.\*  
31: em\_hlg\_inv:.\*  
32: em\_hlg\_other:.\*  
33: em\_hlg\_mus:.\*  
34: em\_hlg\_pln:.\*  
35: em\_hlg\_rod:.\*  
36: em\_hlg\_man:.\*  
37: em\_hlg\_vrl:.\*  
38: em\_sy:.\*  
39: em\_hlg\_hum:.\*  
40: em\_hlg\_mus:.\*  
41: em\_hlg\_other:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	959	100.0	798	10 BC008261	BC008261 Mus muscu
2	959	100.0	1017	10 AF186115	AF186115 Mus muscu
3	959	100.0	158198	2 AC012013	AC012013 Homo sapi
4	938	97.8	814	6 AX440456	AX440456 Sequence
5	938	97.8	814	9 AB015631	AB015631 Homo sapi
6	938	97.8	1080	9 AY032624	AY032624 Homo sapi
7	938	97.8	1210	6 AX464040	AX464040 Homo sapi
8	474.5	49.5	155023	2 AC025574	AC025574 Homo sapi
9	472	49.2	182350	2 AC073896	AC073896 Homo sapi
10	472	49.2	184762	2 AC023500	AC023500 Homo sapi
11	451.5	47.1	207424	10 AC090489	AC090489 Genomic s
12	429	44.7	101666	2 AC109891	AC109891 Rattus no
13	361	37.6	649	9 AF186113	AF186113 Homo sapi
14	361	37.6	824	9 BC001027	BC001027 Homo sapi
15	336.5	35.1	69753	2 AC103156	AC103156 Rattus no
16	262	27.3	564	6 AC040375	AC040375 Sequence
17	248	25.9	2378	9 AK094445	AK094445 Homo sapi
18	212.5	22.2	996	3 AY094698	AY094698 Drosophi
19	193.5	20.2	40824	3 U13070	U13070 Caenorhabdi
20	159	16.6	785	8 AY088187	AY088187 Arabidops
21	156.5	16.3	622	5 AB055671	AB055671 Dario rer
22	145	15.1	175963	3 AC009379	AC009379 Drosophi
23	145	15.1	279530	3 AE003518	AE003518 Drosophi
24	142.5	14.9	1286	3 AY060987	AY060987 Drosophi
25	140	14.6	55359	2 AC012854	AC012854 Drosophi
26	140	14.6	172904	3 AC007414	AC007414 Drosophi
27	140	14.6	275390	3 AE003831	AE003831 Drosophi
28	136.5	14.2	744	6 AR142811	AR142811 Sequence
29	136.5	14.2	746	6 AX136616	AX136616 Sequence
30	136.5	14.2	747	6 BD006701	BD006701 Novel pol
31	136.5	14.2	1413	9 BC019903	BC019903 Homo sapi
32	136.5	14.2	1419	6 AX136423	AX136423 Sequence
33	136.5	14.2	1470	6 AX376412	AX376412 Sequence
34	136.5	14.2	1470	6 AX454716	AX454716 Sequence
35	136.5	14.2	1470	6 AX491194	AX491194 Sequence
36	136.5	14.2	1512	6 AR142810	AR142810 Sequence
37	136.5	14.2	1547	6 BC033339	BC033339 Sequence
38	133	13.9	161566	9 AC008060	AC008060 Homo sapi
39	127.5	13.3	76187	2 AC015219	AC015219 Homo sapi
40	123	12.8	168418	2 AC009403	AC009403 Homo sapi
41	116	12.1	1260	6 AX380553	AX380553 Sequence
42	116	12.1	181896	2 AC114126	AC114126 Rattus no
43	116	12.1	201649	2 AC120813	AC120813 Rattus no
44	115.5	12.0	201649	2 AC120813	AC120813 Rattus no
45	113	11.8	210631	2 AC129184	AC129184 Mus muscu

RESULT 1

## ALIGNMENTS

BC008261 LOCUS BC008261 798 bp mRNA linear ROD 07-AUG-2002  
 DEFINITION Mus musculus, transmembrane protein 4, clone MGC:6853  
 IMAGE:2650612, mRNA, complete cds.  
 ACCESSION BC008261  
 VERSION BC008261.1 GI:14198400  
 KEYWORDS MGC.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 798)  
 REFERENCE Direct Submission  
 TITLE Submitted (22-May-2001) National Institutes of Health, Mammalian  
 JOURNAL Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 REMARK NHI-MGC Project URL: <http://mgc.ncl.nih.gov>  
 COMMENT Contact: MGC help desk  
 Email: [gcgaps-rt@mail.nih.gov](mailto:gcgaps-rt@mail.nih.gov)  
 Tissue Procurement: Lohar Heminghaus Ph.D., Chu-Xia Deng Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,  
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
 Richards, S., Gibbs, R.A.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 5 Row: b Column: 23  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA g1: 9903606.  
 FEATURES  
 source  
 1. 798  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /map="129,C57BL/6J,FVB/N"  
 /clone="MGC:6853 IMAGE:2650612"  
 /tissue\_type="Mammary tumor. Brca1-/-;1; MTV-Cre model. 10  
 months old, gross tissue."  
 /clone\_id="NCI\_CGAP\_Mam3"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 115. 663  
 /codon\_start=-1  
 /product="transmembrane protein 4"  
 /protein\_id="AA08261.1"  
 /db\_xref="GI:14198401"  
 /db\_xref="LOCUSID:56530"  
 /translation="MKGWGLALLGLVLTANARRSODLHCACRLVDELEWIKR  
 VDEKTIQSGFRINPDGSSVVEVPAREAHTELEVCDDMKRGQIDPSTR  
 KNVYRVSRGSESELDLOGIRIDISGTLKFACEISIVEYEDELIEFSSREADNVK  
 DKLSKRTDLCDAHLHRSDEL"  
 BASE COUNT 227 a 182 c 228 g 161 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,29e-88 Length: 798  
 Score: 959.00 Matches: 182  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-082-502-19 (1-182) x BC008261 (1-798)  
 QY 1 MetylsGlyTTPGlyTTPleuAlaLeuLeuGlyValLeuLeuGlyTThAlaTTPAla 20

|||||  
 Db 115 ATGAAGGCTGGGCTTGGCTACCTCTACTTGGGGGCTCTGCGGAACGCGGCGCT 174  
 QY 21 ArgArSerGlnAspLeuHiscysGlyAlaCysAsgAlaLeuValAspGluLeuGlyTTP 40  
 Db 175 CGAAGGACCAAGATCTACACTGTGAGCTTCACAGGCTCTGTGTGATGAAATTAAGTGG 234  
 QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 Db 235 GAAATGGCCCGGTGGAGCCCAAGAACATTCAGATGGGATCTCTCCGATCAATCCA 294  
 QY 61 AspGlySerGlnSerValValGluValProtyrAlaArgSerGluAlaHisLeuThrGlu 80  
 Db 295 GATGGACCCAGTCAGTGTGGAGGTACCTTAATGCCCTCAGAGCCACCTCCACAGAG 354  
 QY 81 LeuLeuGluGluValAlaCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100  
 Db 355 TTGCTTGAAGAGAGTGTGTGACCGAATGAAGAGTACGGGGAACAGATTAACCTTCTAC 414  
 QY 101 HisArgLysAsnTyrValAlaArgValValSerArgAsnGlyGluSerGluLeuAspLeu 120  
 Db 415 CACCCCAAGAACTACGTACGCGCTCGTAGCCGGAATGGAATCCAGTAACTAGACTTA 474  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 Db 475 CAGGCAATCCGAATGACTCAGATATCAGCGCACCCCTCAAGTTTGCCTGTAGAGCAATT 534  
 QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
 Db 535 GTGGAAGATAGAGAGATGAGACTTAATCAATCTCTCCAGAGAGCTGACACACTTAA 594  
 QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisasp 180  
 Db 595 GACAAACTTGGAGTAAAGCGGACAGATCTATGTACCATGCGCTGCACAGATCTCAGAT 654  
 QY 181 GluLeu 182  
 Db 655 GAGCTA 660  
 RESULT 2  
 AF186115 1017 bp mRNA linear ROD 13-JAN-2000  
 LOCUS AF186115  
 DEFINITION Mus musculus putative secreted protein ZS199 (Zs199) mRNA, complete  
 cds.  
 ACCESSION AF186115 GI:6014635  
 VERSION AF186115.1  
 KEYWORDS  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1017)  
 REFERENCE Sheppard, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J. and  
 O'Hara, P.  
 TITLE Mus musculus putative secreted protein  
 JOURNAL Unpublished  
 2 (bases 1 to 1017)  
 AUTHORS Sheppard, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J. and  
 O'Hara, P.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-SEP-1999) Biomolecular Informatics, ZymoGenetics,  
 Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA  
 FEATURES  
 source  
 1. 1017  
 /organism="Mus musculus"  
 /db\_xref="dbEST:AA049839"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:476061"  
 1. 1017  
 /gene="Zs199"  
 343. 891  
 /gene="Zs199"  
 /codon\_start=-1

/product="putative secreted protein ZS19"  
 /protein\_id="AA01433.1"  
 /db\_xref="gi:6014636"  
 /translation="MKGMWLLALIGLSTAMARRSODLHGACRAVDELEMEIAR  
 VPKRTIOMSPFRINPGOSVPEVPRSAHLELEEDCDEKREKEDIDSTHR  
 KNYVAVRNGESSSELDIQRIRISDISIGTLKFACESTIVEEYEDLEIFFSREADNVK  
 DKLSKRDLCDHAIHRSHDEL"  
 BASE COUNT 248 a 250 c 315 g 204 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 4.44e-88 Length: 1017  
 Score: 959.00 Matches: 182  
 Percent Similarity: 100.00% Conserved: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-10-082-502-19 (1-182) x AF186115 (1-1017)

QY 1 MetLysGlyTTPGLYTPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTTPAla 20  
 DB ATGAAGGCTGGGGTGGCTGACCTACTTTGGGGGCTCTGGTGAAGTGGCTGGCT 402  
 QY 21 ArgArgSerGlnAspLeuHisCysGlyValCysArgAlaLeuValAspGluLeuGlu 40  
 DB CGAAGACCCAGATCTACACTGTGAGCTGACAGGCTCTGGTGAATTTGAGAGTGG 462  
 QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 60  
 DB GAAATTCGCCCGCTGACCCCAAGAACCACTGATGGATGCTTCCTCCAGATCAATCCA 522  
 QY 61 AspGlySerGlnSerValAlaGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 80  
 DB GATGGACCCAGTCTGCTGTGAGTACTTATGCCCTCAGAGGCCACCTCAGAGAG 582  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGluThrGlyGluGluGluLeuAspProSerThr 100  
 DB TTGCTTGAGAGGTGTGTGACCGAATGAGAGTACGGGAAACGATTTGACCTTCTTAC 642  
 QY 101 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
 DB CACCGAAGAACTACGATACGCGCTGTGACCGAATGAGAGTCACTGAACTGACTTGA 702  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB CAGGCGATCCGAATTTGACTGATATCAGCGGACCTCTCAAGTTGCTGTGAGAGCATTT 762  
 QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
 DB GTGGAAAGAAATACGAGATGAGTATGAAATCTTCTCCAGAGAGGCTGACAAAGTTAA 822  
 QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 DB GACAAACTTTGCAATGACGCGACAGATCTATGTGACCATGCTCCACAGATCTCAGAT 882  
 QY 181 GluLeu 182  
 DB 883 GAGCTA 888

RESULT 3  
 AC012013 158198 bp DNA linear HTG 17-JUL-2001  
 LOCUS  
 DEFINITION Homo sapiens chromosome 12 clone RP11-764L14, WORKING DRAFT  
 AC012013  
 VERSION AC012013.18 GI:14717292  
 KEYWORDS HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFTN.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 158198)

## AUTHORS

Muzny D.M., Adams C., Adio-Oduola B., Ali-Usman F.R., Allen C.,  
 Alsbrooks S.L., Amaratunga H.C., Are J.R., Banks T., Barbara J.,  
 Benton J., Bimaga K., Blankenburg K., Bonlin D., Bouck J.,  
 Bowle S., Brileva M., Brown E., Brown M., Bryant N.P., Buha C.,  
 Burch P., Burkett C., Burrell K.L., Byrd N.C., Caron T.F.,  
 Carter M., Cavazos S.R., Chacko J., Chavez D., Chen G., Chen R.,  
 Chen Z., Chowdhury I., Christopoulos C., Cleveland C.D., Cox C.,  
 Coyle M.D., Dethorne S.R., David R., Davila M.L., Davis C.,  
 Davy-Carroll L., Dederich D.A., Delaney K.R., Delgado O.,  
 Duggan-Rocha S., Durbin K.J., Earnhart C., Edgar D., Edwards C.C.,  
 Elhaj C., Escotto M., Falls T., Ferraguto D., Flagg N., Ford J.,  
 Foster P., Frantz P., Gabisi A., Gao J., Garcia A., Garner T.,  
 Garza N., Gill R., Gotrell J.H., Guetara M., Gunaratne P., Hale S.,  
 Hamilton K., Harris C., Harris R., Hart M., Havlak P., Hayes A.,  
 Hernandez J., Hernandez O., Hodgson A., Hognes M., Holloway C.,  
 Hollins B., Homsl F., Howard S., Huber J., Huliy S., Hume J.,  
 Jackson L.E., Jacobson B., Jia Y., Johnson R., Jolivet S.,  
 Joudah S., Karlsson E., Kelly S., Khan U., King L., Korvah J.,  
 Kovar C., Kratovic J., Kresh A., Landry N., Leal B., Lewis L.C.,  
 Lewis L., Li J., Li Z., Lichthage O., Lieu C., Liu J., Liu W.,  
 Louieged H., Lozano R.J., Lu X., Lucier A., Lucier R., Luna R.,  
 Ma J., Maheshwari M., Mapua P., Martin R., Martindale A.,  
 Martinez E., Massey E., Mawlin E., McLeod M.P., Meador M.,  
 Mel G., Metzker M., Miner G., Miner Z., Mitchell T., Mohabbat K.,  
 Morgan M., Morris S., Moser M., Neal D., Newton J., Newton N.,  
 Nguyen A., Nguyen N., Nguyen N., Nickerson E., Nwokwuo S.,  
 Ogun M., Okunoye G., Oragunye N., Oviedo R., Pace A., Payton B.,  
 Peery J., Perez L., Peters L., Pichens R., Primus E., Pu L.L.,  
 Quiles M., Ren Y., Rives M., Rojas A., Rojupokan I., Rolfe M.,  
 Ruiz S., Savary G., Scherer S., Scott G., Shen H., Shooshari N.,  
 Sisson I., Sodergren E., Sonalle T., Sparks A., Stanley H.,  
 Stone H., Sutton A., Svatek A., Tabor P., Tamerisa A., Tamerisa K.,  
 Tang H., Tansey J., Taylor C., Taylor V., Telford B., Thomas R.,  
 Thomas S., Usmani K., Vasquez L., Vera V., Villalob D., Vinson R.,  
 Wall R., Wang S., Ward-Moore S., Warren R., Washington C.,  
 Watlington S., Williams G., Williams A., Wlezyn R., Wooden S.,  
 Worley K., Wu C., Wu Y.F., Zhou J., Zorilla S., Nelson D.,  
 Weinstein G. and Gibbs R.

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Direct Submission  
 2 (bases 1 to 158198)  
 Worley, K.C.  
 Submitted (19-OCT-1999) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 13, 2001 this sequence version replaced gi.14547727.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information  
 Center project name: HMKZ

Center clone name: RP11-764L14

----- Summary Statistics  
 Sequencing vector: M13; L08821

Assembly program: Phrap; version 0.990329

Consensus quality: 161278 bases at least 40

Consensus quality: 170806 bases at least 30

Consensus quality: 174843 bases at least 20

Estimated insert size: 168803; sum-of-contigs estimation

Estimated insert size: 163288; agarose-gel estimation

Quality coverage: 5.8x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
 NOTE: This is a 'working draft' sequence. It currently  
 consists of 5 contigs. The true order of the pieces  
 is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 108229: contig of 108229 bp in length  
\* 108230 108329: gap of unknown length  
\* 108330 150246: contig of 41917 bp in length  
\* 150247 150346: gap of unknown length  
\* 150347 153410: contig of 3064 bp in length  
\* 153411 153510: gap of unknown length  
\* 153511 155739: contig of 2229 bp in length  
\* 155740 155839: gap of unknown length  
\* 155840 158198: contig of 2359 bp in length.

FEATURES  
Source  
1. 158198  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-764L14"

BASE COUNT 46057 a 30922 c 31747 g 49030 t 442 others  
ORIGIN

## Alignment Scores:

Pred. No.: 2.26e-85 Length: 158198  
Score: 959.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-10-082-502-19 (1-182) x AC012013 (1-158198)

OY 1 MetLysGlyTyrPglYTrPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
Db 152111 ATGAAGAGCTGGGCTGGCTAGCCCTACTTTGGGGCTGCTGGAACTGGGCT 152170  
OY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
Db 152171 CGAAGGACCAAGATCTACCTGTGAGCTGACAGGCTGTGGTGAATTAAGATGG 152230  
OY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
Db 152231 GAAATGGCCCGCTGAGCCCAAGAACACATTCAGATGGATCTTCCGAATCAATCCA 152290  
OY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
Db 152291 GATGGCACCCAGTCAGTTGTGAGGTACTTATGCCCTCAGAGGCCACCTCCACAGAG 152350  
OY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
Db 152351 TTGCTGGAGAGGTGTGTGACCAATGAAGATAGGGGAACAAATTGACCTTCTACG 152410  
OY 101 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
Db 152411 CACGGCAAGAACTACGTACGGGTGTGAGCCGGAATGAGAAATCCAGTAAGTACTTA 152470  
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 152471 CAGGCAATCCCAATTCAGATCATCAGCCGACCCCTCAAGTTTGCCTGTGAGAGCAT 152530  
OY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
Db 152531 GTGGAGAAATACGAGGATGACCTTTCGAATCTTCTCCAGAGGCTGACAAAGCTTAA 152590  
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
Db 152591 GACAAACTTTCAGATGAAGCGACAGATCATATGATCCATGCCCTGCACAGATTCACAGAT 152650  
OY 181 GluLeu 182  
Db 152651 GACCTA 152656

RESULT 4  
AX440456  
LOCUS AX440456 814 bp DNA linear PAT 28-JUN-2002  
DEFINITION Sequence 309 from Patent WO0190154.  
ACCESSION AX440456  
VERSION AX440456.1 GI:21665266  
KEYWORDS  
SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 Xu, J., Mitcham, J.L., Harlocker, S.L., Dillon, D.C., Secrist, H.,  
Lodes, M.D., Algate, P.A., Fling, S.P., Mannion, V., Benson, D.R. and  
Carter, D.

TITLE  
Compositions and methods for the therapy and diagnosis of ovarian  
cancer

JOURNAL  
Patent: WO 0190154-A 309 29-NOV-2001;

FEATURES  
CORIXA CORPORATION (US)  
Location/Qualifiers  
1. 814  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 210 a 194 c 241 g 169 t  
ORIGIN

## Alignment Scores:

Pred. No.: 4.54e-86 Length: 814  
Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: 6 Gaps: 0

US-10-082-502-19 (1-182) x AX440456 (1-814)

OY 1 MetLysGlyTyrPglYTrPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
Db 145 ATGAAGAGCTGGGCTGGCTAGCCCTACTTTGGGGCTGCTGGAACTGGGCT 204  
OY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
Db 205 CGAGAGACCCAGATCTCAGCTGTGAGATGACAGGCTGTGGTGAATTAAGATGG 264  
OY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
Db 265 GAAATGGCCCGCTGAGCCCAAGAACACATTCAGATGGATCTTCCGAATCAATCCA 324  
OY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
Db 325 GATGGCACCCAGTCAGTTGTGAGGTACTTATGCCCTCAGAGGCCACCTCCACAGAG 384  
OY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
Db 385 CTGCTGGAGAGATATGTGACCGCATGAAAGATATGGGGAACACATTCCTTCCACC 444  
OY 101 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
Db 445 CATGCAAGAACTACGTACGGGTGTGAGCCGGAATGAGAAATCCAGTAAGTACTTA 504  
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 505 CAGGCAATCCCAATTCAGATCATCAGCCGACCCCTCAAGTTTGCCTGTGAGAGCAT 564  
OY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
Db 565 GTGGAGAAATACGAGGATGACCTTTCGAATCTTCTCCAGAGGCTGACAAAGCTTAA 624  
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
Db 625 GACAAACTTTCAGATGAAGCGACAGATCATATGATCCATGCCCTGCACAGATTCACAGAT 684  
OY 181 GluLeu 182

DB	685	GAGCTA	690
RESULT 5			
LOCUS	AB015631		
DEFINITION	AB015631	814 bp	mRNA linear
ACCESSION	AB015631.1		PRI 26-JUL-2001
VERSION	AB015631.1		
KEYWORDS	type II membrane protein.		
SOURCE	Homo sapiens gastric adenocarcinoma CDNA to mRNA, clone, l1p:pkAL-meta-1 clone:HP10390.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 Yokoyama-Kobayashi, M., Yamaguchi, T., Sekine, S. and Kato, S.		
TITLE	Selection of cDNAs encoding putative type II membrane proteins on the cell surface from a human full-length cDNA bank		
JOURNAL	Gene 228 (1-2), 161-167 (1999)		
MEDLINE	99173880		
REFERENCE	2 (bases 1 to 814)		
AUTHORS	Kato, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JUN-1999)		
	Seishi Kato, Research Institute of National		

FEATURES	Location/Qualifiers
source	1. .814

CDS

BASE COUNT	210 a	194 c	241 g	169 t
ORIGIN				

Alignment Scores:	
Pred. No.:	4,54e-86
Score:	938.00
Percent Similarity:	98.3%
Best Local Similarity:	97.2%
Query Match:	97.81%
DB:	9
Length:	814
Matches:	177
Conservative:	3
Mismatches:	2
Indels:	0
Gaps:	0

05-10-082-502-19 (1-182) X AB015631 (1-814)

QY	1	Metysylslyrpyllyrpreuialaleuenuenglyvalleuenuenglythralatrpala	20
Db	145	ATGAAAGCGTGGGGTTCGCTGGCCCTGCTTGGGGCCCTCGTGGGAACCGCTGGGT	20
QY	21	ArgsrsercslasplseuhtsacsllyalacysargalaleuValaspjlugenlurp	40
Db	205	CGAGAGCCAGGATCTCCACTGTGTGACCAAGCGGCTCTGTGTGATACATCAAAATGG	26
QY	41	GlutllealargValaspProlylslythrlllelunmetlyserphearglleasnpro	60
Db	265	GAAATTCGCCAGTGTGACCCCAAGAACCATTCAGATGGGATCTTCCGGATCAATCCA	32
QY	61	AspGlsrclnserValalgluValproctyrAlarsercsluAlhtsleuthrglu	80

Dd	325	GATGGCAGCAGCACTCACTGTGTGGAGGTGCCCTTAATGCCCTAGAGAGGCCACCTCACAGAG	384
Qy	81	LeuleuGluGluValCysAspArgMetLysGluArgGluGluGlnIleAspProSerThr	100
Dd	385	CTGCTGGAGAGATATGTGACCGGATGAAAGAGAGTGTGGGAAACAGATTGATCTTCCACC	444
Qy	101	HisArgLysAsnTyrValArgValValSerArgAsnGlyIuSerSerGluLeuAspLeu	120
Dd	445	CATGCCAAGAACTACGTACGTGTAGTGGCCGGAAATGGAGAAATCCAGTCAACTGAGACCTA	504
Qy	121	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle	140
Dd	505	CAAGCATCCGAGATCGACTCAGATTATTAACCGGCAACCTTCAAGTTTGGCTGGAGACATT	564
Qy	141	ValGluGluTyrArgIuAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys	160
Dd	565	GTCGAGGAATACGAGAGATGAACTATTGCAATCTTTTCCCGAGAGGGGTGCATATGTTTAA	624
Qy	161	AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisLysArgSerHisAsp	180
Dd	625	GACAAACTTTCACATAAGCGAAACAGATCTTTGTGACCATGCCCTGCACAAATTCGATGAT	684
Qy	181	GluLeu 182	
Dd	685	GAGCTA 690	

## RESULT 6

DEFINITION	Homo sapiens saposin-like protein mRNA, complete cds.	2000 bp	linear	PRJ 18-APR-2002
ACCESSION	AY032624			

SOURCE	ORGANISM
Homo sapiens.	Homo sapiens

1 (bases 1 to 1080)  
Bornhauser, B. C., Olsson, P.-A. and Lindholm, D.  
NSAP is a novel saposin-like protein that interacts with MIR and stimulates neurite outgrowth

AUTHORS  
 TITLE  
 JOURNAL  
 DIRECT SUBMISSION  
 SUBMITTED (13-APR-2001) Neuroscience, Uppsala University  
 Husargatan 3, Uppsala 75123, Sweden  
 Received

source	location/qualifiers
	1. .1080

CDS

```

CDS
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q15"
419..967
/note="NSAP; encodes a type 2 membrane signal and a
putative ER retention motif"
/codon_start=1
/product="saposin-like protein"
/protein_id="AAK38148.1"
/db_xref="GI:20196199"
/translation="MKSGMGLALLGLGTPAMARSDILHCGACRALVDELLENIAC
VDPKRTIOMGSFRINPDGOSVAVPYARSAAHLLTELLEIICDNRKEGEQIDISTNRH
KNRYRVGRNGSESELDLOGIRINDSIDGLTKFACESIVEEYEDLEIEFFSREADNWVK
DKLCSKRTDLCDAHLATSHDEL"
238 a 269 c 343 g 230 t
BASE COUNT
ORIGIN

```

Alignment Scores:	
Pred. No.:	6.44e-86
Score:	938.00
Percent Similarity:	98.33%
Best Local Similarity:	97.25%
Query Match:	97.81%
Length:	1080
Matches:	177
Conservative:	2
Mismatches:	3
Indels:	0

DB: 9 Gaps: 0

US-10-082-502-19 (1-182) x AY032624 (1-1080)

OY 1 MettysgltYrPglYTrPleuAlaleuLeuGlyValleuLeuGlyThrAlaTrpAla 20  
|||||  
DB 419 ATGAAGGCTGGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 478

OY 21 ArgArSergInAspLeuHiscysAlaAcysArgAlaLeuValAspIuLeuGluTrp 40  
|||||  
DB 479 CGAGAGACAGACAGATCTCAGCTGAGACATGACAGGCTCTGCTGAGATGAAATAGG 538

OY 41 GluIleAlaArgValAspProLysIsthrIleGlnMetGlySerPheArgIleAsnPro 60  
|||||  
DB 539 GAAATGGCCAGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 598

OY 61 AspGlySergInSerValAlaGluValProTyrAlaArgSergIuAlaHiscysThrGlu 80  
|||||  
DB 599 GATGGCAGCAGTCACTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 658

OY 81 LeuLeuGluGluValCysAspArgMetLysGlyIsthrGluGlnIleAspProSertThr 100  
|||||  
DB 659 CTGCTGGAGAGATATGTGACCGAGATGAGAGATGAGAGATGAGATGATGATCCAC 718

OY 101 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyIsthrSergIuLeuAspLeu 120  
|||||  
DB 719 CATCCCAAGAACTAGTACGTGTAGTGGCGGAGATGAGATCCAGTGAATGATCCAC 778

OY 121 GlnGlyIleArgIleAspSertAspIleSergIsthrLeuLysPheAlaCysGluSertIle 140  
|||||  
DB 779 CAGGCAATCCGAATGACTCAATATTTAGCGGACCTCAAGTTTGGCTGTAGAGCAT 838

OY 141 ValGluIsthrGluAspGluLeuIleGluPheSertArgGluAlaAspAsnValLys 160  
|||||  
DB 839 GTGGAGAAATCGAGAGAACTCATTAATCTTTCCCGAGAGCTGACAAATGTAAA 898

OY 161 AspLysLeuCysSertLysArgThrAspLeuCysAspHiscysAlaLeuHiscysSertIsthr 180  
|||||  
DB 899 GACAAACTTTGACATGACGAACAGATCTTTGTGACCATGCCCTGCACATATGCATGAT 958

OY 181 GluLeu 182  
|||||  
DB 959 GAGCTA 964

RESULT 7  
AX464040 1210 bp DNA linear PAT 16-JUL-2002  
LOCUS Sequence 173 from Patent WO0140466.  
DEFINITION AX464040  
ACCESSION AX464040  
VERSION AX464040.1 GI:21899037  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 Baker,K.P., Beresini,M., DeForge,L., Desnoyers,L., Filvaroff,E.,  
Gao,W.O., Gerritsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,  
Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,  
Wood,W.L. and Zhang,Z.  
SECRETED and transmembrane polypeptides and nucleic acids encoding  
TITLE  
JOURNAL Patent: WO 0140466-A 173 07-JUN-2001;  
Genentech Inc. (US)  
FEATURES  
source 1. 1210 Location/Qualifiers  
BASE COUNT 261 a 316 c 379 g 254 t  
ORIGIN  
Alignment Scores: 7.41e-86 Length: 1210

Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: 6 Gaps: 0

US-10-082-502-19 (1-182) x AX464040 (1-1210)

OY 1 MettysgltYrPglYTrPleuAlaleuLeuGlyValleuLeuGlyThrAlaTrpAla 20  
|||||  
DB 543 ATGAAGGCTGGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 602

OY 21 ArgArSergInAspLeuHiscysAlaAcysArgAlaLeuValAspIuLeuGluTrp 40  
|||||  
DB 603 CGAGAGACAGACAGATCTCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 662

OY 41 GluIleAlaArgValAspProLysIsthrIleGlnMetGlySerPheArgIleAsnPro 60  
|||||  
DB 663 GAAATGGCCAGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 722

OY 61 AspGlySergInSerValAlaGluValProTyrAlaArgSergIuAlaHiscysThrGlu 80  
|||||  
DB 723 GATGGCAGCAGTCACTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 782

OY 81 LeuLeuGluGluValCysAspArgMetLysGlyIsthrGluGlnIleAspProSertThr 100  
|||||  
DB 783 CTGCTGGAGAGATATGTGACCGAGATGAGAGATGAGAGATGAGATGATGATCCAC 842

OY 101 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyIsthrSergIuLeuAspLeu 120  
|||||  
DB 843 CATCCCAAGAACTAGTACGTGTAGTGGCGGAGATGAGATCCAGTGAATGATCCAC 902

OY 121 GlnGlyIleArgIleAspSertAspIleSergIsthrLeuLysPheAlaCysGluSertIle 140  
|||||  
DB 903 CAGGCAATCCGAATGACTCAATATTTAGCGGACCTCAAGTTTGGCTGTAGAGCAT 962

OY 141 ValGluIsthrGluAspGluLeuIleGluPheSertArgGluAlaAspAsnValLys 160  
|||||  
DB 963 GTGGAGAAATCGAGAGAACTCATTAATCTTTCCCGAGAGCTGACAAATGTAAA 1022

OY 161 AspLysLeuCysSertLysArgThrAspLeuCysAspHiscysAlaLeuHiscysSertIsthr 180  
|||||  
DB 1023 GACAAACTTTGACATGACGAACAGATCTTTGTGACCATGCCCTGCACATATGCATGAT 1082

OY 181 GluLeu 182  
|||||  
DB 1083 GAGCTA 1088

RESULT 8  
AC025574 155023 bp DNA linear HTG 09-MAY-2002  
LOCUS AC025574/c  
DEFINITION Homo sapiens chromosome 12 clone RP11-348M3, WORKING DRAFT  
SEQUENCE, 12 unordered pieces.  
ACCESSION AC025574  
VERSION AC025574.13 GI:20428723  
KEYWORDS HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 155023)  
Munzy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaralungu,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barbarta,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowle,S., Brieve,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
David,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Kocha,S., Durbin,K.J.,  
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,







Pred. No.: 6,35e-36 Length: 182350  
Score: 472.00 Matches: 98  
Percent Similarity: 65.13% Conservative: 1  
Best Local Similarity: 64.47% Mismatches: 1  
Query Match: 49.22% Indels: 52  
DB: 2 Gaps: 1

US-10-082-502-19 (1-182) x AC073896 (1-182350)

69 ValProTyrAlaArgSerGluAlaHisLeuThrGluLeuLeuGluGluValCysAspArg 88  
Db 172027 GTGCGCTATGCCCCGCTCAGAGCCGCTCAGAGAGCTGCTGAGAGATATGACCG 171968  
89 MetLysGluTyrGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 108  
Db 171967 ATGAAGAGATATGAGGAGAGATATGATCTTCCACCATGCGAAGATACGATAGTGA 171908  
109 ValSerArgAspGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 128  
Db 171907 GTGGCGCGGAGATGAGATATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 171848  
129 IleSerGlyThrLeuLysPheAla 136  
Db 171847 ATTAGCGGACACCTCAAGTTTGGGGTACCTATGAGGATCGTACCTGTTCTTGACAT 171788  
136  
Db 171787 TAAGGGGTGTTGAGAGACCAATTAGAAAGTATGCTGATATCTTCTCTCTTTGG 171728  
136  
136  
Db 171727 AAGTGAGGAGAGAGCGCTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 171668  
137 CysGluSerIleValGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 156  
Db 171667 TGTGAGAGCATTTGTGAGAGATATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 171608  
157 AspAsnValLysAspLysLeuGluGluGluGluGluGluGluGluGluGluGluGlu 168  
Db 171607 GACAAATGTTAAAGCAAACTTTCAGTAAGCA 171572  
RESULT 10  
LOCUS AC023500 184762 bp DNA linear HTG 24-AUG-2002  
DEFINITION Homo sapiens chromosome 12 clone RP11-183H16, WORKING DRAFT  
AC023500  
AC023500  
AC023500 32 GI:22095070  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULFILLTOP.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 184762)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Albrooks,S.L., Amaratunga,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,H.H.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,J., Gantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorelli,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Hawlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgeson,A., Hogues,M., Holloway,C., Hollins,B.,  
HomsI,F., Howard,S., Huber,J., Huliyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H.,  
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapue,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Maxwell,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,N., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokwesi,S., Ogih,M., Okunou,G.,  
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peary,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rivers,A., Rojas,A., RojudoKan,I., Rolfe,M., Ruiz,S., Savary,G.,  
Scherer,S., Scott,G., Shen,H., Shooshbari,N., Sisson,I.,  
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,R., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D.,  
Wenstock,G., and Gibbs,R.  
Direct Submission  
Unpublished  
2 (bases 1 to 184762)  
Worley,K.C.  
Direct Submission  
Submitted (15-FEB-2000) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 184762)  
Worley,K.C.  
Direct Submission  
Submitted (24-AUG-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Aug 4, 2002 this sequence version replaced gi:22094227.  
Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Drafting Center Code: BCM  
Contact: hgsc-help@bcm.tmc.edu  
Project Information  
Center project name: HAPC  
Center clone name: RP11-183H16  
Summary Statistics  
Sequencing vector: plasmid;  
Sequencing vector: M13;  
Chemistry: Dye-terminator Big Dye; 98% of reads  
Chemistry: Dye-terminator Big Dye; 98% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 183845 bases at least Q40  
Consensus quality: 184849 bases at least Q40  
Consensus quality: 183447 bases at least Q20  
Estimated insert size: 159026; sum-of-contigs estimation  
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation  
NOTE: Estimated insert size may differ from sequence length  
(see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
NOTE: This is a 'working draft' sequence. It currently  
consists of 9 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.  
1 2040: contig of 2040 bp in length  
\* 2041 2140: gap of unknown length  
\* 2141 4170: contig of 2030 bp in length  
\* 4171 4270: gap of unknown length  
\* 4271 8294: contig of 4024 bp in length  
\* 8295 8394: gap of unknown length

\* 8395 14196: contig of 5802 bp in length  
 \* 14197 14296: gap of unknown length  
 \* 14297 33076: contig of 18780 bp in length  
 \* 33077 33176: gap of unknown length  
 \* 33177 52295: contig of 19119 bp in length  
 \* 52296 52395: gap of unknown length  
 \* 52396 84185: contig of 31790 bp in length  
 \* 84186 84285: gap of unknown length  
 \* 84286 120023: contig of 35738 bp in length  
 \* 120024 120123: gap of unknown length  
 \* 120124 184762: contig of 64639 bp in length.

FEATURES  
 SOURCE  
 1. 184762  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /clone="RP11-183H16"

BASE COUNT 47425 a 44751 c 45101 g 46673 t 812 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 6,46e-36 Length: 184762  
 Score: 472.00 Matches: 98  
 Percent Similarity: 65.13% Conservative: 1  
 Best Local Similarity: 64.47% Mismatches: 1  
 Query Match: 49.22% Indels: 52  
 Gaps: 2  
 Gaps: 1

US-10-082-502-19 (1-182) x AC023500 (1-184762)

Oy 69 ValProTyrAlaArgSerGluAlaHisLeuThrGluLeuLeuGluValCysAspArg 88  
 Db 43486 GTGCTTATGCCCCGTCCAGAGCCACCTCACAGAGCTGCGAGAGATATGACCGG 43545  
 Oy 89 MetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValArgVal 108  
 Db 43546 ATGAAGAGTATGGGGAACAAGTTATCTTCACACCATCGAAGAACTACTAGCTGTA 43605  
 Oy 109 ValSerArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 128  
 Db 43606 GTGGCCCGGAATGAGAAATCCAGTCACTGACCTCAACAAGGCATCGAATCGACTCAGAT 43665  
 Oy 129 IleSerGlyThrIleuLysPheAla----- 136  
 Db 43666 ATTACGGCACCTCAAGCTTGGCGGTGAGCTATGGGAATCGGTAGCTGCTTGGACAT 43725  
 Oy 136 ----- 136  
 Db 43726 TAAAGGCTTTGGAGAACCAATTAGAAAGTTAGGCTGATATCTTGTCCCTCTTGG 43785  
 Oy 136 ----- 136  
 Db 43786 AAGTGAAGCAAGACCTCTTATTGCCCTGTGTCACCCAACTTTCCTGGGTGGCAG 43845  
 Oy 137 CysGluSerIleValGluGlyTyrGluAspGluLeuIleGluPhePheSerArgGluAla 156  
 Db 43846 TGTGAGACATTTGTGAGGAATACAGATGACATCACTCAATCTTTTCCGAGAGGCT 43905  
 Oy 157 AspAsnValLysAspLysLeuGlySerLysArgThr 168  
 Db 43906 GACATATGTTAAAGACAACTTTCGAGTAAAGGACAA 43941

RESULT 11  
 AC090489 207424 bp DNA linear ROD 26-JAN-2002  
 LOCUS AC090489  
 DEFINITION Genomic sequence for Mus musculus, clone RP23-104010, complete  
 sequence.  
 AC090489  
 VERSION AC090489.8 GI:18376849  
 KEYWORDS HTG.  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

# REFERENCE

1 (bases 1 to 207424)  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

# AUTHORS

McCombie, W.R., de la Bastide, M., Spiegel, L., Preston, R.,  
 Kirchoff, K., Kuit, R., Nascimento, L., Zuber, T., Ballja, V.,  
 Bell, M., Baker, J., Santos, L., Miller, B., Katzenberger, F., and  
 Muller, S., King, L., Yang, C., Palmer, L., O'Shaughnessy, A. and  
 Dedhia, N.

# TITLE

Genomic sequence for Mus musculus, clone RP23-104010, complete

# JOURNAL

Unpublished

# REFERENCE

2 (bases 1 to 207424)

# AUTHORS

McCombie, W.R.

# TITLE

Direct Submission

# JOURNAL

Submitted (24-FEB-2001) Lila Annenberg Hazen Genome Sequencing  
 Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
 Harbor, NY 11724, USA

# REFERENCE

3 (bases 1 to 207424)

# AUTHORS

McCombie, W.R.

# TITLE

Direct Submission

# JOURNAL

Submitted (26-JAN-2002) Lila Annenberg Hazen Genome Sequencing  
 Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
 Harbor, NY 11724, USA

# COMMENT

On Jan 26, 2002 this sequence version replaced gi:18201765.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30);  
 an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest.

# FEATURES

Location/Qualifiers

# SOURCE

1. 207424  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="RP23-104010"  
 /clone\_lib="RPCT-23"  
 48132..48165

# misc\_feature

/note="We believe the assembly to be correct. The sequence  
 is a mononucleotide (T) repeat in which the exact number  
 of Ts is unknown. Other subclones in the region show one  
 fewer T than that represented in the assembly."

BASE COUNT 52613 a 48671 c 49663 g 56477 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 8.92e-34 Length: 207424  
 Score: 451.50 Matches: 100  
 Percent Similarity: 49.26% Conservative: 0  
 Best Local Similarity: 49.26% Mismatches: 0  
 Query Match: 47.08% Indels: 103  
 Gaps: 1

US-10-082-502-19 (1-182) x AC090489 (1-207424)

Oy 69 ValProTyrAlaArgSerGluAlaHisLeuThrGluLeuLeuGluValCysAspArg 88  
 Db 75789 GTACCTTATGCCCCGTCCAGAGCCACCTCACAGAGTGTGAGAGCTGTGACCGA 75848  
 Oy 89 MetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValArgVal 108  
 Db 75849 ATGAAGAGTATGGGGAACAAGTTATCTTCACACCATCGAAGAACTACTAGCGCTC 75908  
 Oy 109 ValSerArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 128  
 Db 75909 GTGAGCCGGAATGAGAAATCCAGTCACTTACAGGCAATCCGAATGACTCAGAT 75968  
 Oy 129 IleSerGlyThrIleuLysPheAla----- 136  
 Db 75969 ATCAGCGCACCTCAAGATTTGGGTGAGTTACCGCTGTGCGGTGAGTACTGTTCTTG 76028  
 Oy 136 ----- 136  
 Db 76029 GGAGTTATAGGATTCCTGGCACAGATGGAGATGATGCGAATGTTGCTTAC 76088



20238 21736: contig of 1499 bp in length  
21737 21836: gap of unknown length  
21837 23619: contig of 1783 bp in length  
23620 23719: gap of unknown length  
23720 25907: contig of 2188 bp in length  
25908 26007: gap of unknown length  
26008 27406: contig of 1399 bp in length  
27407 27506: gap of unknown length  
27507 28837: contig of 1331 bp in length  
28838 28938: gap of unknown length  
28939 30229: contig of 1292 bp in length  
30230 30329: gap of unknown length  
30330 32012: contig of 1683 bp in length  
32013 32112: gap of unknown length  
32113 33198: contig of 1086 bp in length  
33199 33298: gap of unknown length  
33299 35519: contig of 2221 bp in length  
35520 35619: gap of unknown length  
35620 37677: contig of 2058 bp in length  
37678 37777: gap of unknown length  
37778 38937: contig of 1159 bp in length  
38937 39036: gap of unknown length  
39037 40273: contig of 1237 bp in length  
40274 40373: gap of unknown length  
40374 41906: contig of 1533 bp in length  
41907 42006: gap of unknown length  
42007 44506: contig of 2500 bp in length  
44507 44606: gap of unknown length  
44607 46803: contig of 2197 bp in length  
46804 46903: gap of unknown length  
46904 49226: contig of 2323 bp in length  
49227 49326: gap of unknown length  
49327 51433: contig of 2107 bp in length  
51434 51533: gap of unknown length  
51534 53763: contig of 2230 bp in length  
53764 53863: gap of unknown length  
53864 56613: contig of 2750 bp in length  
56614 56713: gap of unknown length  
56714 58773: contig of 2060 bp in length  
58774 58873: gap of unknown length  
58874 61527: contig of 2654 bp in length  
61528 61627: gap of unknown length  
61629 63192: contig of 1565 bp in length  
63193 63292: gap of unknown length  
63293 65362: contig of 2070 bp in length  
65363 65462: gap of unknown length  
65463 67438: contig of 1976 bp in length  
67439 67538: gap of unknown length  
67539 70505: contig of 2967 bp in length  
70506 70605: gap of unknown length  
70606 74524: contig of 3919 bp in length  
74525 74624: gap of unknown length  
74625 76737: contig of 2113 bp in length  
76738 76837: gap of unknown length  
76839 80133: contig of 3296 bp in length  
80134 80233: gap of unknown length  
80234 83122: contig of 2889 bp in length  
83123 83222: gap of unknown length  
83223 86731: contig of 3509 bp in length  
86732 86831: gap of unknown length  
86833 89929: contig of 3098 bp in length  
89930 90029: gap of unknown length  
90030 93107: contig of 3078 bp in length  
93108 93207: gap of unknown length  
93208 97419: contig of 4212 bp in length  
97420 97519: gap of unknown length  
97520 101666: contig of 4147 bp in length.

Location/Qualifiers  
1. 101666

FEATURES  
Source  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-319E6"

BASE COUNT 24650 a 22934 c 23569 g 23059 t 7454 others

## ORIGIN

Alignment Scores: 7.06e-32 Length: 101666  
Pred. No.: 97 Matches: 97  
Score: 429.00  
Percent Similarity: 49.25%  
Best Local Similarity: 48.74% Mismatches: 2  
Query Match: 44.73% Indels: 99  
DB: 2 Gaps: 1

US-10-082-502-19 (1-182) x AC109891 (1-101666)

QY 69 ValProTyrAlaArgSerGluAlaHisLeuThrGluLeuGluGluValCysAspArg 88  
DB 97572 GTTCTTATGCCCCCTCAGAGCCCTCACGAGGTTGCTTACGAGATATGACCGA 97631  
QY 89 MetLysGluTyrGluGluGlnLeuAspProSerThrHisArgLysAspTyrValArgVal 108  
DB 97632 ATGAGGAGTATGGGAAACAGTATGACCTTCTACCCACCGCAACACTACCTCGTGT 97691  
QY 109 ValSerArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 128  
DB 97692 GTCGGCCGGAACGAGAGATCCAGTACACTACAGGCGTATCCGATTCAGAT 97751  
QY 129 IleSerGlyThrLeuLysPhe----- 135  
DB 97752 ATCACTGCGACCCCTCAAGTTGCGGTGATATGCGTGCAGTACGTGTTCTTG 97811  
QY 135 ----- 135  
DB 97812 GGAGTTCAATAGGATTCCTGGGCACAGATTGGAGTCCGGCGGATATGTTGTCATG 97871  
QY 135 ----- 135  
DB 97872 CTTTGAACCGGCGAGTGAGAGAGCAGTGTTGAAGATTGGCACTAGTTACTAGT 97931  
QY 135 ----- 135  
DB 97932 TCTAGGCGACGTGAGCATATCAATAAACCTGCTATAAATCCAGCTGCCACCA 97991  
QY 135 ----- 135  
DB 97992 AAACGAGGAGAGACTGAGCCGCTGCTCTGCCCCGCTGAGATCCTGGTCACTTCT 98051  
QY 136 -----Ala-CysGluSerIleValGluGluTyrGluAspGluLeuIleG1 150  
DB 98052 CTGTCTTCTGTCGCGCAGTGTGAGACATTTGTGAGAGATGAGATGAACCTTATGA 98111  
QY 150 uPhePheSerArgGluAlaAspAsnValLysAspLysLeuGlySerLysArgThr 168  
DB 98112 ATCTTTTCAAGAGAGGCTGACCAACGTTAAAGACAACCTTGCACTAAGCGGACA 98166

RESULT 13  
AF186113 649 bp mRNA linear PRI 13-JAN-2000  
LOCUS AF186113 Homo sapiens putative secreted protein ZS1G9 (ZS1G9) mRNA, complete cds.  
DEFINITION AF186113 Homo sapiens putative secreted protein ZS1G9 (ZS1G9) mRNA, complete cds.  
ACCESSION AF186113  
VERSION AF186113.1 GI:6014631  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 649)  
AUTHORS Sheppard,P., Jelinek,L., Whitmore,T., Blumberg,H., Lehner,J. and O'Hara,P.  
TITLE Homo sapiens putative secreted protein  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 649)  
AUTHORS Sheppard,P., Jelinek,L., Whitmore,T., Blumberg,H., Lehner,J. and O'Hara,P.  
TITLE Direct Submission



SOURCE

Norway rat.

Rattus norvegicus

Consensus quality: 15506 bases at least Q40  
Consensus quality: 17125 bases at least Q30  
Consensus quality: 18271 bases at least Q20

REFERENCE  
AUTHORS

1 (bases 1 to 69753)  
Munry,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Albrooks,S.L., Amaratunga,H.C., Are,J.R., Ayalew,M., Banks,T., Barbata,J., Benton,J., Blamege,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowle,S., Brileva,M., Brown,M., Brown,M., Bryant,N.P., Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Girelli,D., H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Huliyil,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., Kling,L., Korvah,J., Kovar,C., Kratoch,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichteberg,O., Liew,C., Liu,J., Liu,W., Louised,H., Lozano,R., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okunolu,G., Oregunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Reyes,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sison,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,S., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleciyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,U., Zorrilla,S., Nelson,D., Welnsstock,G. and Gibbs,R.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL

Unpublished  
2 (bases 1 to 69753)  
Worley,K.C.

REFERENCE  
AUTHORS  
JOURNAL

Submitted (24-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 69753)

REFERENCE  
AUTHORS  
JOURNAL

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jul 11, 2002 this sequence version replaced gi:17974643.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GJIX

Center clone name: CH230-208H17

Summary Statistics

Sequencing: Dye-terminator Big Dye 100k of reads

Assembly program: Phrap, version 0.990329

NOTE: Estimated insert size may differ from sequence length.  
NOTE: This is a working draft sequence. It currently consists of 42 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	1366:	contig of 1366 bp in length
1367	1466:	gap of unknown length
1467	2632:	contig of 1166 bp in length
2633	2732:	gap of unknown length
2733	3919:	contig of 1187 bp in length
3920	4019:	gap of unknown length
4020	5169:	contig of 1150 bp in length
5170	5269:	gap of unknown length
5270	6813:	contig of 1544 bp in length
6814	6913:	gap of unknown length
6914	8119:	contig of 1206 bp in length
8120	8219:	gap of unknown length
8220	9384:	contig of 1165 bp in length
9385	9484:	gap of unknown length
9485	10996:	contig of 1512 bp in length
10997	11096:	gap of unknown length
11097	12554:	contig of 1458 bp in length
12555	12654:	gap of unknown length
12655	14376:	contig of 1722 bp in length
14377	14476:	gap of unknown length
14477	15790:	contig of 1314 bp in length
15791	15890:	gap of unknown length
15891	17010:	contig of 1120 bp in length
17011	17110:	gap of unknown length
17111	18308:	contig of 1198 bp in length
18309	18408:	gap of unknown length
18409	19786:	contig of 1378 bp in length
19787	19886:	gap of unknown length
19887	21145:	contig of 1259 bp in length
21146	21245:	gap of unknown length
21246	22539:	contig of 1294 bp in length
22540	22639:	gap of unknown length
22640	23861:	contig of 1222 bp in length
23862	23961:	gap of unknown length
23962	25395:	contig of 1434 bp in length
25396	25495:	gap of unknown length
25496	26515:	contig of 1020 bp in length
26516	26615:	gap of unknown length
26616	28329:	contig of 1714 bp in length
28330	28429:	gap of unknown length
28430	30046:	contig of 1617 bp in length
30047	30146:	gap of unknown length
30147	31613:	contig of 1467 bp in length
31614	31713:	gap of unknown length
31714	33120:	contig of 1407 bp in length
33121	33220:	gap of unknown length
33221	34615:	contig of 1395 bp in length
34616	34715:	gap of unknown length
34716	35867:	contig of 1152 bp in length
35868	35967:	gap of unknown length
35968	37671:	contig of 1704 bp in length
37672	37771:	gap of unknown length
37772	39476:	contig of 1705 bp in length
39477	39576:	gap of unknown length
39577	41221:	contig of 1645 bp in length
41222	41321:	gap of unknown length
41322	43130:	contig of 1809 bp in length
43131	43230:	gap of unknown length
43231	44734:	contig of 1504 bp in length







GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:24:04 ; Search time 1535.7 Seconds  
(without alignments)  
3070.029 Million cell updates/sec

Title: US-10-082-502-20

Perfect score: 848  
Sequence: 1 RRSODLHCACRAVLDELEW.....LCKRPTDLDHALISHDEL 162

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+p2n.model -DEV=xlp  
-O=/gen2.1/USPTO\_spool/US10082502/runat.24032003.135059.6547/app.query.fasta.1.1308  
-DB=GenEmbl -QFWT=fastap -SUFFIX=range -MINMATCH=0.1 -IOFCTL=0 -IOFEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=tbl -NOM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10082502 -ECGN.1.1.4758 -runat.24032003.135059.6547 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_JMAP -LARGEQUER -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pin:\*  
35: em\_htg\_fod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	848	100.0	814	6	AX440456
2	848	100.0	814	9	AB015631
3	848	100.0	1080	9	AY032624
4	848	100.0	1210	6	AX464040
5	830	97.9	798	10	BC008261
6	830	97.9	1017	10	AF186115
7	830	97.9	158198	2	AC012013
8	488.5	57.6	155023	2	AC025574
9	481	56.7	182350	2	AC073896
10	481	56.7	184762	2	AC023500
11	446.5	52.7	207424	10	AC090489
12	436	51.4	101666	2	AC109891
13	257	30.3	649	9	AF186113
14	257	29.0	2378	9	BC001027
15	246	26.5	69753	2	AC103156
16	224.5	22.5	996	3	AY094698
17	202	23.8	40824	3	U13070
18	190.5	18.3	622	5	AB055671
19	155.5	18.2	564	6	AY088187
20	154	18.2	785	8	AY088187
21	154	16.7	175963	3	AC009379
22	141.5	16.7	279530	3	AE003518
23	141.5	16.6	55359	2	AC012854
24	141	16.6	172804	3	AC007414
25	141	16.6	275390	3	AE003831
26	141	16.3	1286	3	AY060987
27	138.5	15.4	161566	9	AC008060
28	131	15.2	744	6	AR142811
29	129	15.2	746	6	AX136616
30	129	15.2	747	6	BD006701
31	129	15.2	1413	9	BC019903
32	129	15.2	1419	6	AX136423
33	129	15.2	1470	6	AX376412
34	129	15.2	1470	6	AX454716
35	129	15.2	1470	6	AX491194
36	129	15.2	1512	6	AR142810
37	129	15.2	1547	6	BC032339
38	129	14.6	76187	2	AC015219
39	124	13.7	168418	9	AC009403
40	116	13.6	181896	2	AC114126
41	115	13.6	201649	2	AC120813
42	115	13.5	201649	2	AC120813
43	114.5	13.5	1280	6	AX380553
44	113	13.3	1280	6	AX129184
45	111	13.1	210631	2	AX129184

RESULT 1

ALIGNMENTS

AX440456  
 LOCUS AX440456 814 bp DNA linear PAT 28-JUN-2002  
 DEFINITION Sequence 309 from Patent WO0190154.  
 ACCESSION AX440456  
 VERSION AX440456.1 GI:21665266  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS 1  
 Xu, J., Mitcham, J. L., Harlocker, S. L., Dillon, D. C., Secrist, H.,  
 Lodes, M. J., Algate, P. A., Fling, S. P., Mannion, J., Benson, D. R. and  
 Carter, D.  
 TITLE Compositions and methods for the therapy and diagnosis of ovarian  
 cancer  
 JOURNAL Patent: WO 0190154-A 309 29-NOV-2001;  
 CORIXA CORPORATION (US)  
 FEATURES  
 SOURCE 1. 814  
 Location/Qualifiers  
 BASE COUNT 210 a 194 c 241 g 169 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3.92e-77 Length: 814  
 Score: 848.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-082-502-20 (1-162) x AX440456 (1-814)

OY 1 ArgatgSerGlnAspLeuHnSCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 |||||||  
 DB 205 CGGAGGACCGAGGATCTCCACTGTGGAGCATGACGGCTGTGGATGAACTAGAAATGG 264  
 OY 21 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
 |||||||  
 DB 265 GAAATGGCCAGCTGAGCCCAAGAACCAATTCAGATGGGATCTTCCGATCAATCCA 324  
 OY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuHrgu 60  
 |||||||  
 DB 325 GATGGCAGCCAGTCACTGAGTGGAGTGGCTTATGCCCCGCTCAGAGCCCACTCAGAG 384  
 OY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 |||||||  
 DB 385 CTGCTGGAGGAGATATGTGACCGGATGAAGAGTATGGGGAAACAGATTGATCTCCACC 444  
 OY 81 HisArgLysAsnTyrValAlaArgValAlaIleArgAsnGlyGluSerSerGluLeuAspLeu 100  
 |||||||  
 DB 445 CATCCCAAGAACTACGTACGTACTGTGGCCGGAATGAGAAATCCAGTAACCTGACCTTA 504  
 OY 101 GlnGlyIleArgTyrIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 |||||||  
 DB 505 CAAAGCATTCCGATGATCTCAGATATTAAGCCGACCTCAAGTTGGCTGTGAGAGCATT 564  
 OY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
 |||||||  
 DB 565 GTGGAGGAATAGCAGAGGAACTCATTTGAAATTTCTTCCGAGACGCGACAGATTTAAA 624  
 OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 |||||||  
 DB 625 GACAAACTTTCAGTAAACGAAACAGATCTTTGTGACATGCCCTGCACATATGCGATGAT 684  
 OY 161 GluLeu 162  
 |||||||  
 DB 685 GAGCTA 690  
 RESULT 2  
 AB015631

AB015631  
 LOCUS AB015631 814 bp mRNA linear PRI 26-JUL-2001  
 DEFINITION Homo sapiens mRNA for type II membrane protein, complete cds,  
 clone:HP10390.  
 ACCESSION AB015631  
 VERSION AB015631.1 GI:4586839  
 KEYWORDS type II membrane protein.  
 SOURCE Homo sapiens gastric adenocarcinoma cDNA to mRNA,  
 clone\_H1b:pkal-meta-1 clone:HP10390.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS 1  
 Yokoyama-Kobayashi, M., Yamaguchi, T., Sekine, S. and Kato, S.  
 TITLE Selection of cDNAs encoding putative type II membrane proteins on  
 the cell surface from a human full-length cDNA bank  
 JOURNAL Gene 228 (1-2), 161-167 (1999)  
 MEDLINE 99173880  
 REFERENCE 2 (bases 1 to 814)  
 AUTHORS Kato, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-JUN-1998) Selshi Kato, Research Institute of National  
 Rehabilitation Center for the Disabled, Department of  
 Rehabilitation Engineering, 4-1 Namiki, Tokorozawa, Saitama  
 359-8555, Japan (E-mail: selshi@rehab.go.jp  
 Tel:042-995-3100(ex.2568), Fax:042-995-3132)

FEATURES  
 SOURCE 1. 814  
 Location/Qualifiers  
 CDS  
 1. 814  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="HP10390"  
 /tissue\_type="gastric adenocarcinoma"  
 /clone\_lib="pkal-meta-1"  
 /cdoon\_start=1  
 /product="type II membrane protein"  
 /protein\_id="PAA76498.1"  
 /db\_xref="GI:4586840"  
 /translation="MKKGRIALLGLALGTRARRSODLHCACRALYDELEETIAQ  
 VDPKLTWGSFRINPDSOSVVEVPARSANHTLEELICDMMKRGEDIDPSTIR  
 KNYRVVGRNSESSELDIGIRIDSDISGLTKFACESIVEREDELIEFESREADNVK  
 DKLSKRTDLCDHALHSHDEL"  
 BASE COUNT 210 a 194 c 241 g 169 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3.92e-77 Length: 814  
 Score: 848.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-082-502-20 (1-162) x AB015631 (1-814)

OY 1 ArgatgSerGlnAspLeuHnSCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 |||||||  
 DB 205 CGGAGGACCGAGGATCTCCACTGTGGAGCATGACGGCTGTGGATGAACTAGAAATGG 264  
 OY 21 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
 |||||||  
 DB 265 GAAATGGCCAGCTGAGCCCAAGAACCAATTCAGATGGGATCTTCCGATCAATCCA 324  
 OY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuHrgu 60  
 |||||||  
 DB 325 GATGGCAGCCAGTCACTGAGTGGAGTGGCTTATGCCCCGCTCAGAGCCCACTCAGAG 384  
 OY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 |||||||  
 DB 385 CTGCTGGAGGAGATATGTGACCGGATGAAGAGTATGGGGAAACAGATTGATCTCCACC 444  
 OY 81 HisArgLysAsnTyrValAlaArgValAlaIleArgAsnGlyGluSerSerGluLeuAspLeu 100  
 |||||||  
 DB 445 CATCCCAAGAACTACGTACGTACTGTGGCCGGAATGAGAAATCCAGTAACCTGACCTTA 504

QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlnSerIle 120  
DB 505 CAAGGATCCGAAATCGACATGATATTACGGCAGCCCTCAAGTTCCGCTGAGACATTT 564  
QY 121 ValGluGlyTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
DB 565 GTGGAGATACGAGATGAACTGATGAAATCTTTTCCCGAGAGGCTGCAATGTTTAA 624  
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
DB 625 GACAACTTTGAGATAGGAGACAGATCTTTGTGACCAATCCCTGCACATATGCATGAT 684  
QY 161 GluLeu 162  
DB 685 GAGCTA 690

RESULT 3  
AY032624 1080 bp mRNA linear PRI 18-APR-2002  
LOCUS AY032624  
DEFINITION Homo sapiens saposin-like protein mRNA, complete cds.  
ACCESSION AY032624  
VERSION AY032624.1 GI:20196198  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Carnivora; Homini; Hominidae; Homo.  
REFERENCE  
AUTHORS Bornhauser, B.C., Olsson, P.-A. and Lindholm, D.  
TITLE NSAP is a novel saposin-like protein that interacts with MTR and stimulates neurite outgrowth  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1080)  
AUTHORS Olsson, P.-A. and Lindholm, D.  
TITLE Direct Submission  
JOURNAL Submitted (13-APR-2001) Neuroscience, Uppsala University,  
Husargatan 3, Uppsala 75123, Sweden  
FEATURES  
SOURCE  
1. 1080  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/map="12q13"  
419..967  
/note="NSAP; encodes a type 2 membrane signal and a putative ER retention motif"  
/codon\_start=1  
/product="saposin-like protein"  
/protein\_id="AAK38148.1"  
/db\_xref="GI:20196198"  
/translation="MKGMGLALLGLLGTAMARRSODLHGCACRALVDLEETIAQ  
VDPKTIQMSFRINPDGOSVVEVPYASFAHLLELEICDPRKREGEQIDSTHR  
KNYRVYGRNGSSSELDOGIRIDISGTLKFAESIVEYEDELIEFSEKADNVK  
DKISKRPTDLCHALHSHDEL"

BASE COUNT 238 a 269 c 343 g 230 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5,33e-77 Length: 1080  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-082-502-20 (1-162) x AY032624 (1-1080)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
DB 479 CGGAGAGCCAGCATCTCCACCTGTGGAGATCGACGGCTCTGTGGATGAACTAGAAATGG 538  
QY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40

DB 539 GAAATGGCCAGGTGGACCCCAAGAAACCATTCAGATGGATCTTCCGATCAATCCA 598  
QY 41 AspGlySerGlnSerValValGluValProTyrAlaAspSerGluAlaHisLeuThrGlu 60  
DB 599 GATGGCACCAGCTGCTGCTGTGGAGGTGCTTATGCTCCGCTGAGAGGCCACCTCAGAG 658  
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
DB 659 CTGCTGGAGAGATATGACCCGATGAAGAGATATGGGAAACAGATTCCTTCACCC 718  
QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerGlnLeuAspLeu 100  
DB 719 CATGCCAAGAACTACGTACGTGTAGTGGCCGGAATGAGAAATCACTGAACTGACCTA 778  
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlnSerIle 120  
DB 779 CAAGGATCCGAAATCGACATGATATTACGGCAGCCCTCAAGTTCCGCTGAGACATTT 838  
QY 121 ValGluGlyTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
DB 839 GTGGAGATACGAGATGAGATGAACTGATGAAATCTTTTCCCGAGAGGCTGACATGTTAA 898  
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
DB 899 GACAACTTTGAGATAGGAGACAGATCTTTGTGACCAATCCCTGCACATATGCATGAT 958  
QY 161 GluLeu 162  
DB 959 GAGCTA 964

RESULT 4  
AX464040 1210 bp DNA linear PAT 16-JUL-2002  
LOCUS AX464040  
DEFINITION Sequence 173 from Patent WO0140466.  
ACCESSION AX464040  
VERSION AX464040.1 GI:21899037  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Carnivora; Homini; Hominidae; Homo.  
REFERENCE  
AUTHORS Baker, R.P., Beresini, M., DeForge, L., Desnoyers, L., Flivveroff, E.,  
Gao, M.O., Gerlitsen, M.E., Goddard, A., Godowski, P.J., Gunney, A.L.,  
Sherwood, S., Smith, Y., Stewart, T.A., Tumas, D., Wetlanbe, C.K.,  
Wood, W.L. and Zhang, Z.  
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding  
JOURNAL same  
JOURNAL Patent: WO 0140466-A 173 07-JUN-2001;  
Genentech Inc. (US)  
FEATURES  
SOURCE  
1. 1210  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 261 a 316 c 379 g 254 t  
ORIGIN

Alignment Scores:  
Pred. No.: 6,35e-77 Length: 1210  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-082-502-20 (1-162) x AX464040 (1-1210)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
DB 603 CGGAGAGCCAGCATCTCCACCTGTGGAGATCGACGGCTCTGTGGATGAACTAGAAATGG 662  
QY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40

```

Db 663 GNAATGGCCAGGTGAGCCCAAGAGACCATTCAGATGGGATCTTCCGGATCAATCCA 722
Oy 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
Db 723 GATGACACCAAGTCAAGTGTGGAGGTGCTTATGCCCCGCTGAGAGCCCACTCACAAG 782
Oy 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
Db 783 CTGCTGGAGGAGATATGTGACCGGATGGAAGGATGTGGGAAACAGATGATCTTCCACC 842
Oy 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerGlnLeuAspLeu 100
Db 843 CATGCCAAGAACTACGTCACGTGTAGTGGCCGGAATGGAAGAACTCCAGTAATGGACCTA 902
Oy 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
Db 903 CAGGCAATCCGAATGAGTCAAGATATTAGCGGCACCCCAAGTTGGCTGTGAGAGCAT 962
Oy 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
Db 963 GTGGAGGAATACGAGATGAACTCATTTCAATCTTCCCGAGAGGCTGACAAATGTATAA 1022
Oy 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
Db 1023 GCAAACTTTGACATGACGAGACAGATCTTTGTGACCATGCCCCGACATATGCGATGAT 1082
Oy 161 GluLeu 162
Db 1083 GAGCTA 1088

RESULT 5
BC008261 798 bp mRNA linear ROD 07-AUG-2002
LOCUS BC008261 IMAGE:2650612, mRNA, complete cds.
DEFINITION BC008261
ACCESSION BC008261.1 GI:14398400
VERSION MGC.
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 798)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Lotmar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Guanaraine, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 5 Row: b Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 9903606.
location/Qualifiers
1..798
/organism="Mus musculus"

```

```

CDS
/db_xref="taxon:10090"
/mdb="129, C57BL/6J, FVB/N"
/clone="MGC:6853 IMAGE:2650612"
/tissue_type="Mammary tumor. Brca1-/fl; MMTV-Cre model. 10
months old, gross tissue."
/clone_lib="NCI_CGAP_Mam"
/lab_host="DH10B"
/ncbi_host="Vector: pCMV-SPORT6"
115..663
/codon_start=1
/product="transmembrane protein 4"
/db_xref="GI:14398401"
/db_xref="LocusID:56530"
/translation="MKGWMLALLGLVLTAMARRSODLHCGRALVDELEIAR
VDPKTIOMGSEFRINPDOSOSVPEPARSAHLLELLEVCDDMKKEGEOIDPSTHR
KNYRVVSRNGSESLDQGRIDSDISGTLKPFACESTVEYEDLEIFPSREADNWK
DKLSKRTDLDAHLHSHDEL"
BASE COUNT 227 a 182 c 228 g 161 t
ORIGIN
Alignment Scores:
Pred. No.: 2,56e-75 Length: 798
Score: 830.00 Matches: 158
Percent Similarity: 98.77% Conservative: 2
Best Local Similarity: 97.53% Mismatches: 0
Query Match: 97.88% Indels: 0
Gaps: 0
DB: 10
US-10-082-502-20 (1-162) x BC008261 (1-798)
Oy 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTyrP 20
Db 175 CGAAGGACCCAAAGATCTACACTGTGAGCTGACAGGCTGTGGAGTAATTAAGATGG 234
Oy 21 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40
Db 235 GAAATTTGCCGGGTGGACCCCAAGAGACCATTCAGATGGATGCTTCCGAATCAATCA 294
Oy 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
Db 295 GATGACACCAAGTCAAGTGTGGAGGTACTTATGCCCCCTGAGAGCCCACTCACAAG 354
Oy 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
Db 355 TTGCTTGAAGGAGTGTGTGACCGAATGAAGGATGACGGGAAACAGATGACCTTACCC 414
Oy 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerGlnLeuAspLeu 100
Db 415 CACCGCAAGAACTACGTCAGCGGTGAGCCGGAATGGAAGAACTCCAGTAATGACTTA 474
Oy 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
Db 475 CAGGCAATCCGAATGAGTCAAGATATTAGCGGCACCCCAAGTTGGCTGTGAGAGCAT 534
Oy 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
Db 535 GTGGAGGAATACGAGATGAACTTATCGAATCTTCCACAGAGGCTGACAACTGATAA 594
Oy 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
Db 595 GACAACTTTGACATGACGAGACAGATCTATGTGACATGCCCCGACAGATCTCAGAT 654
Oy 161 GluLeu 162
Db 655 GAGCTA 660

RESULT 6
AF186115 1017 bp mRNA linear ROD 13-JAN-2000
LOCUS AF186115
DEFINITION Mus musculus putative secreted protein Zs1G9 (Zs1g9) mRNA, complete
cds.
ACCESSION AF186115

```

VERSION AF186115.1 GI:6014635  
 KEYWORDS  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 1017)  
 AUTHORS Sheppard, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J. and O'Hara, P.  
 TITLE Mus musculus putative secreted protein  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1017)  
 AUTHORS Sheppard, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J. and O'Hara, P.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-SEP-1999) Biomolecular Informatics, Zymogenetics, Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA  
 FEATURES  
 source location/Qualifiers  
 1..1017  
 /organism="Mus musculus"  
 /db\_xref="dbEST:AA049839"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:476061"  
 1..1017  
 /gene="Zs1g9"  
 343..891  
 /gene="Zs1g9"  
 /codon\_start=1  
 /product="putative secreted protein ZS1g9"  
 /protein\_id="AF01433.1"  
 /db\_xref="GI:6014636"  
 /translation="MKKGMALLLLGLVLTANRRSODLHCACALYDELEMETAR  
 VDPKRTIOMGSEFRINDGSGVVEVYPAVSALTELEECVCKRMEYEDQIDPSTR  
 KNYVRVRSNGSESLDGIQIRIDISIGLTFACRSIVEEYEDLEEFESREADNVK  
 DKTSRTDCAHLRSHDEL  
 BASE COUNT 248 a 250 c 315 g 204 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,44e-75 Length: 1017  
 Score: 830.00 Matches: 158  
 Percent Similarity: 98.77% Conservative: 2  
 Best Local Similarity: 97.53% Mismatches: 2  
 Query Match: 97.88% Indels: 0  
 Gaps: 0  
 US-10-082-502-20 (1-162) x AF186115 (1-1017)  
 QY 1 ArgArgSerGlnAspLeuHISGysGlyAlaCysArgAlaLeuValAspGluLeuGluTyrP 20  
 DB 403 CGAAGGAGCCAGATCTACACTGTGAGCTTGACAGGCGCTCTGATGATGATTAGATGG 462  
 QY 21 GUAUAGAGUValAspProLysLysThrIleGluMetGlySerPheArgGlyLeuAspPro 40  
 DB 463 GAAATTGCCCGGTGACCCCAAGAGACATTTCACATGGATGGATCCCTCCGATTCATTC 522  
 QY 41 AspglySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 DB 523 GATGGAGCCAGTCACTTGTGGAGTACTTATGCCCGCTCAGAGCCACCTCACAGAG 582  
 QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGluIleAspProSerThr 80  
 DB 583 TTGCTTGAGAGAGTGTGTACCGCAATGAGAGGTACGGGGAACAGATTGACCCCTCTACC 642  
 QY 81 H18AglvAsnTyrValArgValValGluValArgAsnGlyGluSerSerGluLeuAspLeu 100  
 DB 643 CACCCCAAGAACTACCTACGCTCTGAGCGGAGATGGAGATTCACAGTACACTTCA 702  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 DB 703 CAGGCGATCCGATTCATGATCAGATATACGCGGACCCCTCAAGTTTGCCTGTGAGAGCAT 762  
 QY 121 ValGluGluTyrGluValAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValVal 140

DB 763 GTGGAAGAAATAGAGAGGAGGACCTTATCGAATTCCTTCACAGAGGCTCGACCACTTTAAA 822  
 QY 141 AsplyLeuGysSerTyrArgThrAspLeuGysAspHisAlaLeuHisIleSerHisAsp 160  
 DB 823 GACAAACCTTTCAGTAAAGCGAGATCTATGTACCATGCCCTCCAGACATCTCAGCAT 882  
 QY 161 GluLeu 162  
 DB 883 GAGCTA 888  
 RESULT 7  
 AC012013  
 LOCUS  
 DEFINITION Homo sapiens chromosome 12 clone RP11-764L14, WORKING DRAFT  
 AC012013  
 AC012013.18 GI:14717292  
 VERSION  
 AC012013  
 KEYWORDS  
 HTG: HTGS\_PHASE1; HTGS\_DNAFT; HTGS\_ACTIVEFIN.  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 158198)  
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,  
 Alshrocks, S.L., Amaratunga, H.C., Are, J.R., Banks, T., Allen, C.,  
 Benton, J., Blum, K., Blankenburg, K., Bonin, D., Bouch, J.,  
 Bowle, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,  
 Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F.,  
 Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,  
 Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C.,  
 Coyle, M.D., Dethorne, S.R., David, R., Davila, M.L., Davis, C.,  
 Davy, Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,  
 Den, A.L., Ding, X., Dinn, H.H., Douthett, C., Edgar, D., Edwards, C.C.,  
 Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,  
 Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,  
 Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,  
 Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunatane, P., Hale, S.,  
 Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hayes, A.,  
 Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C.,  
 Hollins, B., Homs, F., Howard, S., Huber, J., Hui, Y., Hume, J.,  
 Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolyvet, S.,  
 Joudan, S., Karlsson, E., Kelly, S., Khan, U., Kling, L., Korah, J.,  
 Kovari, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,  
 Lewis, L., Li, J., Li, Z., Lichtenberg, O., Lieu, C., Liu, D., Liu, W.,  
 Louisedge, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
 Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A.,  
 Martinez, E., Massey, E., Mashiney, E., McLeod, M.P., Meador, M.,  
 Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,  
 Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,  
 Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nokenkwo, S.,  
 Ogulu, M., Okunolu, G., Oragunye, N., Oyedero, R., Pace, A., Payton, B.,  
 Peery, J., Perez, L., Peters, L., Pickens, R., Prims, E., Pu, L.L.,  
 Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shooshert, N.,  
 Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H.,  
 Stone, H., Sutton, A., Swalek, A., Taber, P., Tamerisa, A., Tamerisa, K.,  
 Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,  
 Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,  
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,  
 Wellington, S., Williams, G., Williamson, A., Wlecyk, R., Woodson, S.,  
 Worley, K., Wu, C., Wu, Y.F., Zhou, J., Zorilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.  
 Direct Submission  
 Unpublished  
 REFERENCE 2 (bases 1 to 158198)  
 Morley, K.C.  
 TITLE  
 JOURNAL  
 Submitted (19-OCT-1999) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 13, 2001 this sequence version replaced gi:14547727.  
 ----- Genome Center

Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 Project Information  
 Center project name: HMK2  
 Center clone name: RP11-764114  
 Summary Statistics  
 Sequencing Vector: M13; L08821  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 161278 bases at least Q40  
 Consensus quality: 170806 bases at least Q30  
 Consensus quality: 174843 bases at least Q20  
 Estimated insert size: 168803; sum-of-contigs estimation  
 Estimated insert size: 163288; agarose-fp estimation  
 Quality coverage: 6x in Q20 bases; agarose-fp estimation  
 Quality coverage: 5.8x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 NOTE: This is a 'working draft' sequence. It currently  
 consists of 5 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1 108229: contig of 108229 bp in length  
 108230 108329: gap of unknown length  
 108330 150246: contig of 41917 bp in length  
 150247 150346: gap of unknown length  
 150347 153410: contig of 3064 bp in length  
 153411 153510: gap of unknown length  
 153511 155739: contig of 2229 bp in length  
 155740 155839: gap of unknown length  
 155840 158198: contig of 2359 bp in length.

FEATURES  
 Location/Qualifiers  
 1..158198  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /clone="RP11-764114"

BASE COUNT 46057 a 30922 c 31747 g 49030 t 442 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,62e-72 Length: 158198  
 Score: 830.00 Matches: 158  
 Percent Similarity: 98.77% Conservative: 2  
 Best Local Similarity: 97.53% Mismatches: 2  
 Query Match: 97.88% Indels: 0  
 Gaps: 0

US-10-082-502-20 (1-162) x AC012013 (1-158198)

QY 1 ArgArgSerGlnAspLeuNHISySGLYALAcysAlrgalaleuValAspLleuLutrp 20  
 |||||  
 Db 152171 CGAAGGAGCCAAAGATCTACACTGTGAGCTGACAGGGCTGTGTGATGAAATGAGATGG 152230  
 |||||

QY 21 GltuLleAGlnValAspProLysLysLthrLleGlnMetGlySerPheArgLleAsnPro 40  
 |||||  
 Db 152231 GAAATTCGCCCGTGGAGCCCAAGAGCCATTCAGATGGGATCCCTCCGATCAATCA 152290  
 |||||

QY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLleuThcGlu 60  
 |||||  
 Db 152291 GATGGCAGCCAGTACGTTGTGAGAGTACCTTAATGCCCGCTACAGGCCACCTCACAGAG 152350  
 |||||

QY 61 LeuLengLugluLleCysAspArgMetLysGluTyrGlyGlnGlnLleAspProSerThr 80  
 |||||  
 Db 152351 TTGCTTAGAGAGGTGTGTGACCAAGATGAAGGAGTACGGGAGAACAGATTGACCTTCATCC 152410  
 |||||

QY 81 HisArgLysAsnTyrValArgValAlaGlyArganglyGluSerSerGluLeuAspLeu 100  
 |||||  
 Db 152411 CACCCCAAGAACTAGCTACGCTCTGAGCCCGAGTGGAGAAATCCAGATCACTTAA 152470  
 |||||

QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLleuLysPheAlaCysGluSerIle 120  
 |||||  
 Db 152471 CAGGGCATCCGAAATGACTACATATCACGGCACCTTCATGATTTGGTGTGAGACAT 152530  
 |||||

QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
 |||||  
 Db 152531 GTGGAAGATACAGAGATGAGCTTATCGAATCTTCTCCAGAGAGCGACAACTTAA 152590  
 |||||

QY 141 AspLysLleCysSerLysArgThrAspLeuCysAspHisAlaLleHisSerHisAsp 160  
 |||||  
 Db 152591 GACAAACTTTCAGTAAAGCGGACAGATCTATGTGACATGCGCTCCACAGATCTACGAT 152650  
 |||||

QY 161 GltuLeu 162  
 |||||  
 Db 152651 GAGCTA 152656

RESULT 8  
 AC025574/c  
 LOCUS  
 DEFINITION Homo sapiens chromosome 12 clone RP11-348M3, WORKING DRAFT  
 AC025574  
 VERSION AC025574.13 GI:20428723  
 KEYWORDS HTG; HTGS; PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 155023)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-olsan,F.R., Allen,C.,  
 Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
 Barbata,J., Benton,S., Birmange,H.C., Birmange,K., Blankenburg,K., Bonnin,D.,  
 Bouck,J., Boyle,S., Brieve,M., Brown,E., Brown,M., Bryant,N.P.,  
 Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,X., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Havlik,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
 Homsl,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Licharge,O., Lien,C., Liu,J., Liu,W., Lonsdale,H.,  
 Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maneshwari,M., Mapua,P., Martin,R., Matlindale,A., Martinez,E.,  
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,N., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokoko,S., Ogund,M., Okwundu,G.,  
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Rives,M., Rojas,A., Rojubenka,I., Rolfe,M., Rutz,S., Savery,G.,  
 Scherer,S., Scott,G., Shen,H., Shoochitari,N., Sisson,I.,  
 Sodergren,E., Sonelke,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Swalek,A., Tabor,P., Tamerisa,K., Tamerisa,K., Tang,H.,  
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wellington,S.,  
 Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G., and Gibbs,R.

TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 155023)  
AUTHORS Worley, K. C.  
TITLE Direct Submission  
JOURNAL Submitted (11-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 155023)  
Worley, K. C.  
REFERENCE Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On May 3, 2002 this sequence version replaced g1:20335511.  
JOURNAL Genome Center  
Center: Baylor College of Medicine

## COMMENT

Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
Project Information  
Center project name: HALV  
Center clone name: RP11-348M3  
Summary Statistics  
Sequencing vector: M13  
Chemistry: Dye-Primer Bodypy: 5% of reads  
Chemistry: Dye-Terminator Big Dye: 95% of reads  
Assembly program: Phrap: version 0.990329  
Consensus quality: 141522 bases at least Q40  
Consensus quality: 147097 bases at least Q40  
Consensus quality: 150442 bases at least Q20  
Estimated insert size: 150706; sum-of-coverage estimation  
Quality coverage: 4x in Q20 bases; sum-of-coverage estimation

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a working draft sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```

1      3033: contig of 3033 bp in length
*      3034      3133: gap of unknown length
*      3134      5971: contig of 2838 bp in length
*      5972      6071: gap of unknown length
*      6072      10269: contig of 4198 bp in length
*      10270      10369: gap of unknown length
*      10370      13009: contig of 2640 bp in length
*      13010      13109: gap of unknown length
*      13110      17573: contig of 4464 bp in length
*      17574      17673: gap of unknown length
*      17674      24650: contig of 6977 bp in length
*      24651      31778: gap of unknown length
*      31779      31878: gap of unknown length
*      31879      42263: contig of 10385 bp in length
*      42264      42363: gap of unknown length
*      42364      51104: contig of 8741 bp in length
*      51105      51204: gap of unknown length
*      51205      68306: contig of 17102 bp in length
*      68307      109481: gap of unknown length
*      68407      109481: contig of 41075 bp in length
*      109482      109581: gap of unknown length
*      109582      155023: contig of 45442 bp in length.

```

## FEATURES

## SOURCE

1. 155023  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/clone="RP11-348M3"

BASE COUNT	4008 a	36207 c	35987 g	41693 t	1128 others
ORIGIN					
Alignment Scores:					
Pred. No.:	7 01e-38	Length:	155023		
Score:	488.50	Matches:	114		
Percent Similarity:	39.72%	Conservative:	0		
Best Local Similarity:	39.72%	Mismatches:	0		
Query Match:	57.61%	Indels:	173		
DB:	2	Gaps:	2		

```

US-10-082-502-20 (1-162) x AC025574 (1-155023)
QY      49  ValProTALAARserGluAlaHisLeuThGluLeuGluGluLeuIleCysAspArg 68
      |||
Db 56120 GTGGCTTATAGCCCGCTCAAGAGGCCACCTCAGACAGACTCTGGAGAGATATGTGACCGG 56061
QY      69  MetLysGluTyrGlyGluGluIleAspProSerThHisArgLysAsnTyrValArgVal 88
      |||
Db 56060 ATGAAAGAGTATAGGCGAAGACAGATTGATCCCTCCACCCATCGCAAGAACTAGACGTGTA 56001
QY      89  ValGlyArgAsnGlyGluSerSerGluLeuAspProSerThHisArgLysAsnTyrValArgVal 108
      |||
Db 56000 GTGGCCGGAATGAGAAATCCAGTGAATGAGACTGACACTCAAGGCAATCGAATCGACTGAT 55941
QY      109  IleSerGlyThrLeuLysPheAla----- 116
      |||
Db 55940 ATTAGGGGACCCCTCAAGTTTGGGGTGAGACTATGGGAATCGGAGCTGTTTGACAT 55881
QY      116 ----- 116
Db 55880 TAAAGCGTTGTTGAGAACCAATTAGAAAGTTAGCTGATATCTTCTCTCTTTGG 55821
QY      116 ----- 116
Db 55820 AAGTGAGGCAAGAGGCCCTTATTGCCCTGTCTCACCATTTCCTGGTGGCAG 55761
QY      117  CysGluSerTLeValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAla 136
      |||
Db 55760 TTGAGAGAGATTGTGAGAGAAATCGAGATGAACTATGATTTCTTCCGAGAGGCT 55701
QY      137  AspAsnValLysAspLysLeuCysSerLysArgTh----- 148
      |||
Db 55700 GACAAATGTTAAAGACAAACTTTGCACATAAGCGAACGTAAGCTGCCCCACTTATCTC 55641
QY      148 ----- 148
Db 55640 CTGTCTAGTACGCTGTGTGAACCTGACATTCCTTACTCACCATACACCCCATCCGA 55581
QY      148 ----- 148
Db 55580 GAAACCTCTGGATTATAGTCAGAGAGCTGTCTTCATTCCCTTGAGGACTTGGG 55521
QY      148 ----- 148
Db 55520 GTTTCCTCCACATTTTCTGTAGCTTTCAGATTGTGTCCCTTGACCCCTGAGATTAACA 55461
QY      148 ----- 148
Db 55460 GGTGAGAGACAGTGGGAGTGTATATCAATTATTAAGCATATATATCTGCAGAAATGCTGCT 55401
QY      148 ----- 148
Db 55400 TGGTGTGTTTGTGTGACAGCTTGCTGGGAAATTTCTTGATACGCGAGATATACGTA 55341
QY      149 ----- 149
Db 55340 TGAAGATGCGCATATCTTAAATGTTTGCATTTGACATTTGACAGATCTTGTGACCATGCCCTG 55281
QY      156  HisIleSerHisAspGluLeu 162
      |||
Db 55280 CACATATGCGATGATGAGCTTA 55260

```





Oy 116 ----- 116  
 Db 171727 AAGTGAAGCAGAGCCCTTATTCGCCCTGTCACCCATTTCCTGGGTGGCAG 171668  
 Oy 117 CCGGUSGJIEVGLGGLTGTyrLuspgLueuilegLuPheserArGLua 136  
 Db 171667 TGTGAGACATTGTGGAGATACGAGATGATGATTCATTCCTCCGAGAGGCT 171608  
 Oy 137 AspaValLysAspLysLeucCysSerLysArgThrRap-LeucCysAspHisAlaLeuH1 156  
 Db 171607 GACATGTGTAAGACAACTTGTGAGTAGAGACAGATAGCTGCCCTTATTCCTC 171548  
 Oy 156 sllsSer 158  
 Db 171547 CTGTTCT 171541  
 RESULT 10 184762 bp DNA 1linear HTG 24-AUG-2002  
 AC023500  
 LOCUS Homo sapiens chromosome 12 clone RP11-183H16, WORKING DRAFT  
 DEFINITION  
 AC023500  
 VERSION AC023500.32 GI:22095070  
 KEYWORDS HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 184762)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., All-oman,F.R., Allen,C.,  
 Alshrocks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T.,  
 Barbieri,J., Benton,J., Bimge,K., Blankenburg,K., Bonin,T.,  
 Bouck,J., Bowie,S., Blevins,M., Brown,E., Brown,M., Bryant,N.P.,  
 Burch,P., Burckell,C., Butrell,K.L., Byrd,N.C.,  
 Carion,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorbett,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
 Homel,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Krtavcic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,  
 Lozdo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,N., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokwenkwo,S., Oguh,M., Okunju,G.,  
 Oragunye,N., Oviedo,R., Pace,A., Payson,B., Peary,J., Perez,L.,  
 Peters,L., Pickens,R., Pimus,E., Pu,L.L., Qulles,M., Ren,Y.,  
 Rivers,R., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,  
 Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,  
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Usmani,K., Vaequez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williamson,A., Wlaczek,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstock,G., and Gibbs,R.  
 TITLE  
 JOURNAL  
 REFERENCE  
 2 (bases 1 to 184762)  
 Unpublished

AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Drafting Center Code: BCM  
 Contact: hgsc-help@bcm.tmc.edu  
 Project Information  
 Center project name: HARC  
 Center clone name: RP11-183H16  
 Summary Statistics  
 Sequencing vector: Plasmid  
 Sequencing vector: M13  
 Chemistry: Dye-terminator Big Dye 988 of reads  
 Chemistry: Dye-terminator Big Dye 988 of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 183845 bases at least Q40  
 Consensus quality: 184849 bases at least Q30  
 Consensus quality: 185447 bases at least Q20  
 Estimated insert size: 159026; sum-of-contigs estimation  
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation  
 NOTE: Estimated insert size may differ from sequence length  
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
 NOTE: This is a 'working draft' sequence. It currently  
 consists of 9 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.  
 1  
 2041 2040: contig of 2040 bp in length  
 2141 2140: gap of unknown length  
 4171 4170: contig of 2030 bp in length  
 4271 4270: gap of unknown length  
 8295 8294: gap of 4024 bp in length  
 8395 8394: gap of unknown length  
 14197 14196: contig of 5802 bp in length  
 14297 14296: gap of unknown length  
 14297 14296: contig of 18780 bp in length  
 33077 33076: gap of unknown length  
 33177 33176: gap of 19119 bp in length  
 52295 52295: contig of 19119 bp in length  
 52396 52395: gap of unknown length  
 84185 84185: contig of 31790 bp in length  
 84285 84285: gap of unknown length  
 120023 120023: contig of 35738 bp in length  
 120024 120023: gap of unknown length  
 120124 120123: contig of 66639 bp in length.  
 Location/Qualifiers  
 1. 184762  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /clone="RP11-183H16"  
 BASE COUNT 47425 a 44751 c 45101 g 46673 t 812 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.01e-37 Length: 184762  
 Score: 481.00 Matches: 103  
 Percent Similarity: 63.80% Conservative: 1

Best Local Similarity: 63.19% Mismatches: 6  
Query Match: 56.72% Indels: 53  
DB: 2 Gaps: 1

US-10-082-502-20 (1-162) x AC023500 (1-184762)

OY 49 ValProtyrAlaArgSerGluAlaHisIleuThrGluLeuLeuGluGluIleCysAspArg 68  
|||||  
DB 43486 GTGGCTTAATGCCCGTCAGAGGCCACCTCACAGAGCTGGAGAGATATGTGACCGG 43545

OY 69 MetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValArgVal 88  
|||||  
DB 43546 ATGAAGAGGATGGGGAACAATGATCTTCACCCATCCCAAGAACTAGCTAGTCTA 43605

OY 89 ValGlyArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 108  
|||||  
DB 43606 GTGGCCCGGAATGGAGATCCAGTCAACTGAGCTACAAAGCATCCGAATCGACTCAGAT 43665

OY 109 IleSerGlyThrLeuLysPheAla----- 116  
|||||  
DB 43666 ATTAGCGGACCCCTCAAGTTTGGCGGTAGCTATGGGAATCGGTGCTTTGGACAT 43725

OY 116 ----- 116  
DB 43726 TAAGGGTTGTGGAGAACCAATTAGAAAGTTAGGCTGATATCTTGTCTCTCTTTGG 43785

OY 116 ----- 116  
DB 43786 AAGTGAAGCAGCAAGCCCTTATATGCCCTGTGTACCCCATTTCTCCCTGGTTGGCAG 43845

OY 117 CysGluSerIleValGluGlnTyrGluAspGluLeuIleGluPhePheSerArgGluAla 136  
|||||  
DB 43846 TGTAGAGCATGTGGAGGAATACGAGATGAACTCATGTGAATCTTTTCCGAGAGGCT 43905

OY 137 AspAsnValLysAspLysLeuCysSerLysArgThrAsp-LeuCysAspHisAlaLeuHis 156  
|||||  
DB 43906 GACAAATTAAAGCAAACTTTCAGTAAGCAAGTAAGCTAGCTGCCCCCATTTATCTC 43965

OY 156 sIleSer 158  
|||  
DB 43966 CTGTCTT 43972

RESULT 11  
AC090489 207424 bp DNA linear ROD 26-JAN-2002  
LOCUS AC090489  
DEFINITION Genomic sequence for Mus musculus, clone RP23-104010, complete  
sequence.  
AC090489  
VERSION AC090489.8 GI:18376849  
KEYWORDS HTG.  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 207424)  
REFERENCE  
AUTHORS McCormbie,W.R., de la Bastide,M., Spiegel,L., Preston,R.,  
Kirchoff,K., Kuit,K., Nascimento,L., Zucavero,T., Ballia,V.,  
Beil,M., Baker,J., Santos,L., Miller,B., Katzenberger,F.,  
Muller,S., King,L., Yang,C., Palmer,L., O'Shaughnessy,A. and  
Dedhia,N.  
TITLE Genomic sequence for Mus musculus, clone RP23-104010, complete  
sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 207424)  
AUTHORS McCormbie,W.R.  
TITLE Direct Submission  
JOURNAL Submitted (24-FEB-2001) Lita Annenberg Hazen Genome Sequencing  
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
Harbor, NY 11724, USA  
3 (bases 1 to 207424)  
REFERENCE  
AUTHORS McCormbie,W.R.  
TITLE Direct Submission  
JOURNAL Submitted (26-JAN-2002) Lita Annenberg Hazen Genome Sequencing

COMMENT  
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
Harbor, NY 11724, USA  
On Jan 26, 2002 this sequence version replaced gi:18201765.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions all sequencing problems, such  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest.

FEATURES  
source  
1..207424  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-104010"  
/clone\_1fb="RPC1-23"  
48132..48165  
misc\_feature  
/note="We believe the assembly to be correct. The sequence  
is a mononucleotide (T) repeat in which the exact number  
of Ts is unknown. Other subclones in the region show one  
fewer T than that represented in the assembly."

BASE COUNT 52613 a 48671 c 49663 g 56477 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.83e-33 Length: 207424  
Score: 446.50 Matches: 98  
Percent Similarity: 48.77% Conservative: 1  
Best Local Similarity: 48.28% Mismatches: 1  
Query Match: 52.65% Indels: 103  
DB: 10 Gaps: 1

US-10-082-502-20 (1-162) x AC090489 (1-207424)

OY 49 ValProtyrAlaArgSerGluAlaHisIleuThrGluLeuLeuGluGluIleCysAspArg 68  
|||||  
DB 75789 GTACCTTAATGCCCGTCAGAGGCCACCTCACAGAGTGTGTGAGAGGTGTGACCGA 75848

OY 69 MetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValArgVal 88  
|||||  
DB 75849 ATGAAGGAGTACGGGGAACAATGACCTTCTACCCACCGCAAGAACTAGCTAGCGCTC 75908

OY 89 ValGlyArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 108  
|||||  
DB 75909 GTGAGCGGGAATGGAGATCAGTGAACCTTACAGGCGCATCCGAATGACTCAGAT 75968

OY 109 IleSerGlyThrLeuLysPheAla----- 116  
|||||  
DB 75969 ATCAGCGGACCCCTCAAGTTTGGCGGTAGTACCGCTGTGCGGTAGTACCTGTTCTTG 76028

OY 116 ----- 116  
DB 76029 GCACTTAATGAGATTCCTCGACAGATGGAGATGATCGGAAGTGTGCTTAC 76088

OY 116 ----- 116  
DB 76089 TTGTAACCTAAGCACTCCAGGGGAGCGTTGAAGATCAGCTTAAGTTGAGCGTGGC 76148

OY 116 ----- 116  
DB 76149 CTAGTTCATAGTAGTTCGAGGCCAGCTGGCAGATATAGTAAAGCCCTGTCTTAAAT 76208

OY 116 ----- 116  
DB 76209 CCCCACCCCAACCAACCAAGTAGAGAGTGAAGCAGTCTTGGCCACGCTTGGGA 76268

OY 117 ----- CysGluSerIleValGluGlnTyrGlu 125  
|||||  
DB 76269 TCTGGGTCAATTCTCTCTTCTTGGTCCGAGGTGAGAGCATTTGGAAGATACGAG 76328

OY 126 AspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLysAspLysLeuCysSer 145  
|||||  
DB 76329 GATGAGCTTATCGAATCTTCTCCAGAGGCTGACCAACGTTAAAGCAAACTTTGCACT 76388





BASE COUNT 175 a 150 c 171 g 149 t 4 others

## Alignment Scores:

Pred. No.: 2,73e-17 Length: 649  
Score: 257.00 Matches: 49  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 30.31% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-20 (1-162) x AF186113 (1-649)

OY 1 ArgArgSerGlnAspLeuHiscysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20

DB 164 CGGAGAGACCCAGATCTCCAGTGTGAGACATGACGGCTGTGTGGATGAACATAAGATGG 223

OY 21 GluIleAlaGlnValAspProlylsThrIleGlnMetGlySerPheArgIleAsnPro 40

DB 224 GAAATTCGCCAGGTGAGACCCCAAGAACATTCAGATGGATCTTTCCGATCAATCCA 283

OY 41 AspGlySerGlnSerValValGluVal 49

DB 284 GATGGCAGCCAGTCAGTGTGGAGGTA 310

RESULT 14 BC001027 824 bp mRNA linear PRI 12-JUL-2001

ACCESSION BC001027 Homo sapiens, transmembrane protein 4, clone MGC:1545

VERSION BC001027.1 GI:12654402

KEYWORDS MGC.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

FEATURES

SOURCE

1. 824

/organism="Homo sapiens"  
/db\_xref="locusID:10330"  
/db\_xref="taxon:9606"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
Series: RAL Plate: 5 Row: P Column: 13  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6014631.

## CDS

BASE COUNT 221 a 194 c 233 g 176 t

OY 1 ArgArgSerGlnAspLeuHiscysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20

DB 310 CGGAGAGACCCAGATCTCCAGTGTGAGACATGACGGCTGTGTGGATGAACATAAGATGG 369

OY 21 GluIleAlaGlnValAspProlylsThrIleGlnMetGlySerPheArgIleAsnPro 40

DB 370 GAAATTCGCCAGGTGAGACCCCAAGAACATTCAGATGGATCTTTCCGATCAATCCA 429

OY 41 AspGlySerGlnSerValValGluVal 49

DB 430 GATGGCAGCCAGTCAGTGTGGAGGTA 456

RESULT 15 AK094445 2378 bp mRNA linear PRI 15-JUL-2002

ACCESSION AK094445 Homo sapiens CDNA FLJ37126 fls, clone BRACE2022693, weakly similar to Transmembrane protein 4.

VERSION AK094445.1 GI:21753508

KEYWORDS Oligo capping; fls (full insert sequence).

SOURCE Homo sapiens cerebellum CDNA to mRNA, clone\_11b:BRACE2

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE

JOURNAL

REMARK

COMMENT

FEATURES

SOURCE

1. 824

/organism="Homo sapiens"  
/db\_xref="locusID:10330"  
/db\_xref="taxon:9606"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
Series: RAL Plate: 5 Row: P Column: 13  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6014631.

Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7  
Kazusa-Kamata, Kisarazu, Chiba 292-0812, Japan  
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEDO human CDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan. CDNA full insert sequencing: Research Association for Biotechnology (RAB) (CDNA library Key Technology Center etc.); 5 - 6 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and

FEATURES RAB; annotation: HRI and RAB.  
Location/Qualifiers  
Source 1..2378

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="BRACE2022693"  
/tissue\_type="cerebellum"  
/clone\_id="BRACE2"  
/note="cloning vector: PME18SFL3"

BASE COUNT 723 a 432 c 477 g 746 t  
ORIGIN

Alignment Scores:

Pred. No.:	1.73e-15	Length:	2378
Score:	246.00	Matches:	49
Percent Similarity:	62.30%	Conservative:	27
Best Local Similarity:	40.16%	Mismatches:	44
Query Match:	29.01%	Indels:	2
DB:	9	Gaps:	1

US-10-082-502-20 (1-162) x AK094445 (1-2378)

```

QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
   |||||
Db 80 GACGGACTGACTACTACAGGACGACAGATCCCTACCTCAGTCGAGCGCTTCTACCGGAT 139
   |||||

QY 61 LeuLeuGluGluIleCysAspArgMetIleGluTyrGlyGluGlnIleAspProSerThr 80
   |||||
Db 140 CTTTGGAGAAAGCTGTGAGCGAATGACGAACTACAAAGCTTGAAGAAAGACCTGTGAGC 199
   |||||

QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGluSerSerGluLeuAspLeu 100
   |||||
Db 200 AAGGAGAGAACTTTCAGAGATTGCTCTCTAGGAAAGGAGACAAATATACCAAGAAATT 259
   |||||

QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
   |||||
Db 260 AAAAATTTGATTTTATTTCTGATGCTTACAGACCTTGAATTTGCGTGTGAACATATA 319
   |||||

QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
   |||||
Db 320 ATAGAGAGATGATGAAATGAAATCTCTACTTATCGCCCGAGAGACACACTATCTAGCT 379
   |||||

QY 141 AspLysLeuGlySerLysArgThrAspLeuGlyAspHisAlaLeuHisIleSerHisAsp 160
   |||||
Db 380 GACAAAGCTGTGACGTGAATAATCAGATCTGTGAAACTTCT-----GCTAATCATACT 433
   |||||

QY 161 GluLeu 162
   |||||
Db 434 GAGCTC 439

```

Search completed: March 31, 2003, 03:42:06  
Job time : 1675.7 secs

1:	em_estbna.*
2:	em_esthna.*
3:	em_estlin.*
4:	em_estmu.*
5:	em_estov.*
6:	em_estpl.*
7:	em_estro.*
8:	em_hnc.*
9:	gb_est1.*
10:	gb_est2.*
11:	gb_hnc.*
12:	gb_est3.*
13:	gb_est4.*
14:	gb_est5.*
15:	em_estlin.*
16:	em_estom.*
17:	gb_gss.*
18:	em_gss_hmv.*
19:	em_gss_hum.*
20:	em_gss_pln.*
21:	em_gss_vrt.*
22:	em_gss_fun.*
23:	em_gss_man.*
24:	em_gss_mus.*
25:	em_gss_other.*
26:	em_gss_pro.*
27:	em_gss_trod.*

Result No.	Score	Query Match	Length	DB	ID	Description
1	848	100.0	588	14	BM832828	BM832828 K-ESTO107
2	848	100.0	591	14	BM832888	BM832888 K-ESTO107
3	848	100.0	611	13	B1712899	B1712899 I498I09..x
4	848	100.0	673	14	B0668530	B0668530 AGENCOURT
5	848	100.0	697	12	BF344334	BF344334 602014700E
6	848	100.0	719	12	BE733900	BE733900 601568377
7	848	100.0	742	12	BE733900	BE733900 601568377
8	848	100.0	745	12	BE791763	BE791763 602728455
9	848	100.0	930	14	BQ437660	BQ437660 601581588
10	848	100.0	1117	13	BM552907	BM552907 AGENCOURT
11	843	99.4	599	13	B1791489	B1791489 I498I09..x
12	842	99.3	880	14	BQ421827	BQ421827 AGENCOURT
13	840	99.1	783	13	B1550436	B1550436 603192905
14	835	98.5	868	14	BQ421069	BQ421069 AGENCOURT
15	833	98.2	791	12	BG704443	BG704443 602686333
16	830	97.9	665	10	AM919569	AM919569 EST350873
17	830	97.9	674	14	BQ202181	BQ202181 UI-R-DY1-
18	830	97.9	724	14	BQ210416	BQ210416 UI-R-DY1-
19	830	97.9	754	14	AK013568	AK013568 MMS MMSCU
20	830	97.9	772	11	AK007914	AK007914 MMS MMSCU
21	830	97.9	808	13	B1408525	B1408525 602964995
22	830	97.9	843	13	B1412539	B1412539 602990995
23	830	97.9	853	11	AK013014	AK013014 MMS MMSCU
24	830	97.9	1255	11	AK019927	AK019927 MMS MMSCU
25	827	97.5	773	13	B1550162	B1550162 603193045
26	825	97.3	657	14	BQ211867	BQ211867 UI-R-DY1-
27	822	96.9	818	12	BE870315	BE870315 601447464
28	822	96.9	875	13	B1763950	B1763950 603049829
29	821	96.8	887	12	B1762759	B1762759 603048045
30	821	96.5	802	12	BF608664	BF608664 MY1..00160
31	817	96.3	678	10	BE309953	BE309953 601081903
32	817	96.3	891	13	B1408381	B1408381 602964013
33	816	96.2	709	13	B1411532	B1411532 602965126
34	815	96.1	549	10	BE108865	BE108865 ug83d11..x
35	815	96.1	593	10	AM167673	AM167673 ug84e01..x
36	815	96.1	720	13	B1909759	B1909759 603047987
37	809	95.4	797	13	B1762927	B1762927 603047987
38	808	95.3	783	13	B1763784	B1763784 603049729
39	808	95.3	914	12	BE799495	BE799495 601589289
40	807	95.2	746	12	B1551339	B1551339 603193144
41	799	94.2	765	13	BF312821	BF312821 601988804
42	794	93.6	475	10	AM414257	AM414257 u096C08..Y
43	794	93.6	856	12	BF965088	BF965088 602268856
44	792	93.4	595	10	AM414039	AM414039 u098G03..x
45	792	93.4	856	13	B1454180	B1454180 603170602

TITLE Kim Y.S.  
JOURNAL 21C Frontier Korean EST Project 2001  
COMMENT Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 11 row: B column: 03  
High quality sequence stop: 588.  
Location/Qualifiers  
1. 588  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="SSNU484s1-11-B03"  
/clone\_11b="SSNU484s1"  
/sex="M"  
/tissue\_type="Stomach"  
/cell\_type="Epithelial"  
/cell\_line="SNU-484"  
/lab\_host="Top10F"  
/note="Organ: Stomach; Vector: pTZ18RP1; Site:1: EcoRI Site:2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 Promotor as 5' primer and M(DT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F with electroporation method."

BASE COUNT 149 a 140 c 176 g 123 t  
ORIGIN

Alignment Scores:  
Pred. No.: 6,23e-103 Length: 588  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-10-082-502-20 (1-162) x BM832828 (1-588)

QY 1 AAGATGSeGlnAspLeuH1ScysGlyAlaCysArgAlaLeuValAspGluLeuGluTTP 20  
DB 62 CGAGAGAGCCAGATCTCCACCTGTGAGCATGACAGGCTCTGTGATGACATGAAATG 121  
QY 21 GUAUUAAGLValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
DB 122 GAUAATGCCAGGTGAGACCCCAAGACCAATTCATGAGATCTTCCGGATCAATCCA 181  
QY 41 AAspGlySerGlnSerValValGluValProTyrAlaArgSerGlnValIleHisLeuThrGlu 60  
DB 182 GATGGAGCCAGCTCACTGTGTGAGAGTGCCTTATGCCGCTCAGAGGCCACCTCACAGAG 241

QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
DB 242 CTCCTGGAGAGATATCTGACCCGATGAAGAGATAGGGAGACATTCCTCTCAC 301  
QY 81 HisArgLysAspTyrValArgValValGlyArgAsnGlyGlnSerGluLeuAspLeu 100  
DB 302 CATGCGAAGAACTAGCTAGCTGTAGGGCCGGAATGGAAATCCAGTCACTGGACCTA 361  
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlnSerIle 120  
DB 362 CAAGGATCGCATGCATCTCAGATATTAGCGCACCCCTCAAGTTTGGTGTGAGACATT 421  
QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
DB 422 GTGAGAGAAATACAGAGATGAACATCTTATCTTCCGAGAGGCTGCAATGTCTTAA 481  
QY 141 AspLysLeucCysSerLysArgThrAspLeuCysAsnHisAlaLeuHisIleSerHisAsp 160  
DB 482 GACAAACTTTCAGTAAAGCAACAGATCTTGTGACCAATGCCCTGCACATATGCATGAT 541  
QY 161 GluLeu 162  
DB 542 GAGCTA 547

RESULT 2  
BM832888 591 bp mRNA linear EST 06-MAR-2002  
LOCUS K-EST0107445 SSNU484s1 Homo sapiens cDNA clone SSNU484s1-11-H01  
DEFINITION 5', mRNA sequence.  
ACCESSION BM832888  
VERSION BM832888.1 GI:19189297  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 591)  
Kim Y.S., Hahn Y., Oh J.H., Lee J.Y., Ahn H.Y., Chu M.Y., Kim M.R., Oh R.J., Cheong J.E., Sohn H.Y., Kim J.M., Park H.S., Kim S. and Kim Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished (2002)  
CONTACT: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 11 row: B column: 01  
High quality sequence stop: 591.  
Location/Qualifiers  
1. 591  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="SSNU484s1-11-H01"  
/clone\_11b="SSNU484s1"  
/sex="M"  
/tissue\_type="Stomach"  
/cell\_type="Epithelial"  
/cell\_line="SNU-484"  
/lab\_host="Top10F"  
/note="Organ: Stomach; Vector: pTZ18RP1; Site:1: EcoRI; Site:2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of



competent cells E. coli Top10F' by electroporation method. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(OT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F' with electroporation method."

BASE COUNT 150 a 142 c 176 g 123 t  
ORIGIN

## Alignment Scores:

Pred. No.: 6,28e-103 Length: 591  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-10-082-502-20 (1-162) x BM832888 (1-591)

OY 1 ArgArgSerGlnAspLeuHsCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
DB 62 CGAGAGAGCCAGGATCTCCAGTGTGAGCATCCAGGCGCTGTGTGATGACATAGATG 121  
OY 21 GlltleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 40  
DB 122 GAAATGCCAGGTGAGACCCCAAGAACCATTCAGATGGATCTTCCGATCAATCCA 181  
OY 41 AspglySerGlnSerValValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
DB 182 GATGGAGGACCATGATGTGTGAGGTGCTTATGCCCTGACAGGCCACCTCAGAG 241  
OY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
DB 242 CTGCTGGAGGAGATATGTACCGGATGAGAGATGGGGAACAGATTGATCTCCACC 301  
OY 81 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 100  
DB 302 CATCCCAAGACATACGTACGTAGTGGCCGAGATGAGATCCAGTCAATGACCTA 361  
OY 101 GlnGlyIleArgGlnLeaSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 362 CAGGCAATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421  
OY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
DB 422 GTGGAGGAATACGAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 481  
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisLeuSerHisAsp 160  
DB 482 GACAACTTTCAGATGACGAGAACAGATCTTGTGACATCCCTGACATATCCATGAT 541  
OY 161 GluLeu 162  
DB 542 GAGCTA 547

## RESULT 3

LOCUS B1712899 611 bp mRNA linear EST 11-MAR-2002  
DEFINITION id98f09.y1 HR85 islet Homo sapiens cDNA clone IMAGE:5085832.5'

SEQUENCE B1712899  
similar to TR:Q9Y2B0 Q9Y2B0 Type II MEMBRANE PROTEIN. ;, mRNA

ACCESSION B1712899  
VERSION B1712899.1 GI:15688594  
KEYWORDS EST.  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 611)  
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Keestner, K.,  
Lemshka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Hillier, L., Marra, M., Page, D., Wylie, T., Martin, J., Blistain, A.,  
Schmitt, A., Theising, B., Rittler, E., Ronko, I., Bennett, J., Cardenas,  
M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R., Williams, T.,  
Jackson, Y., and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)

## JOURNAL

CONTACT: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biohpc.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@im.wustl.edu)  
Seq primer: -40RP from Glibco  
High quality sequence stop: 463.

## FEATURES

## source

1. 611  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5085832"  
/clone\_1id="HR85 islet"  
/tissue\_type="Purified pancreatic islet"  
/lab\_host="DH10B"  
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1:  
NotI; Site\_2: XhoI; cDNA made by oligo-dT priming.  
Site-selected on agarose gel. Average insert size  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue,  
Metabolism Div. (Alan Pernutt Lab), Washington University,  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."

BASE COUNT 157 a 143 c 184 g 127 t  
ORIGIN

## Alignment Scores:

Pred. No.: 6.6e-103 Length: 611  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-10-082-502-20 (1-162) x B1712899 (1-611)

OY 1 ArgArgSerGlnAspLeuHsCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
DB 90 CGAGAGAGCCAGGATCTCCAGTGTGAGCATCCAGGCGCTGTGTGATGACATAGATG 149  
OY 21 GlltleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 40  
DB 150 GAAATGCCAGGTGAGACCCCAAGAACCATTCAGATGGATCTTCCGATCAATCCA 209  
OY 41 AspglySerGlnSerValValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
DB 210 GATGGAGGACCATGATGTGTGAGGTGCTTATGCCCTGACAGGCCACCTCAGAG 269  
OY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
DB 270 CTGCTGGAGGAGATATGTGACCGGATGAGAGTATGGGGAACAGATTGATGATCTCCACC 329  
OY 81 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 100

```

|||||
Db 330 CATGCCAAGACTAGCTAGCTAGTGGCCGGAATGAGAAATCCAGTGAACCTGGACCTA 389
Oy 101 gInglyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
Db 390 CAAGGATCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 449
Oy 121 ValGluGluThrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
Db 450 GTGGAGAAATACGAGATACATCATGTAATCTTTCCCGAGAGGCTGCATATGTTAAA 509
Oy 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
Db 510 GACAACTTTGCGTAGTAAAGCAACAGATCTTTGACCATGCTCCATGACATATGACATGAT 569
Oy 161 GluLeu 162
Db 570 GAGCTA 575

RESULT 4
B0668530 673 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT 8211118 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6276413
DEFINITION 5', mRNA sequence.
ACCESSION B0668530
VERSION B0668530.1 GI:2178777
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 673)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LDCM2461 row: c column: 06
High quality sequence start: 208
High quality sequence stop: 350.
Location/Qualifiers
1. 673
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6276413"
/clone_lib="NIH-MGC_102"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="FDH10B (phage-resistant)"
/note="Organ: salivary gland; Vector: pORF7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 196 a 149 c 184 g 144 t
ORIGIN

```

```

US-10-082-502-20 (1-162) x B0668530 (1-673)
Oy 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
Db 47 CGGAGACGCCAGATGATCTCCACTGTGAGACATGACAGGCTCTGTGATGACATGATG 106
Oy 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetCysSerPheArgIleAsnPro 40
Db 107 GAATATGCCAGGTGGAGCCCAAGAGAACCATTCAGATGGATCTTCCGATCAATCCA 166
Oy 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
Db 167 GATGGACGCCAGTACGTGGTGGAGGTGCTTATATGCCCTCAAGAGCCCACTCACAGAG 226
Oy 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
Db 227 CTCCTGGAGAGATATGTGACCGGATGAAGAGATATGCGGAACAGATGATCTTCCACC 286
Oy 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyCysSerGluLeuAspLeu 100
Db 287 CATGCCAAGAACTACGTAGCTAGTGGCCGGAATGAGAAATCCAGAACTGGACCTA 346
Oy 101 GInglyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
Db 347 CAAGGATCCGAATCAGATGATATTTAGCGGACCACTCAAGTTTCGTGTGAGACATT 406
Oy 121 ValGluGluThrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
Db 407 GTGGAGAAATACGAGATGACATCATGTAATCTTTCCCGAAGGCTGCATATGTTAAA 466
Oy 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
Db 467 GACAACTTTGCGTAGTAAAGCAACAGATCTTTGACCATGCTCCATGACATATGATGAT 526
Oy 161 GluLeu 162
Db 527 GAGCTA 532

RESULT 5
BF344334 697 bp mRNA linear EST 22-NOV-2000
LOCUS BF344334
DEFINITION 602014708F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4150512
ACCESSION BF344334
VERSION BF344334.1 GI:11291554
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 697)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LAM9413 row: p column: 01
High quality sequence stop: 692.
Location/Qualifiers
1. 697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4150512"
/clone_lib="NCI_CGAP_Brn64"
/tissue_type="gliblastoma with EGFR amplification"

```

```

Alignment Scores:
Pred. No.: 7,62e-103
Score: 848.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Caps: 0

```

/lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.57 kb. Constructed by Life  
 Technologies. Note: this is a NCI-CCAP library."  
 BASE COUNT 173 a 167 c 220 g 137 t  
 ORIGIN

## Alignment Scores:

Pred. No.:	8.02e-103	Length:	697
Score:	848.00	Matches:	162
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-10-082-502-20 (1-162) x BE344334 (1-697)

QY 1 ArgArSergInSArValaIglValaProtyrAlaArgSerGluAlaHisLeuThrGlu 20  
 |||||  
 DB 152 CGGAGGAGCCAGATCTCCACTGTGGAGCATGCCAGGCTGTGGATGAACTAGATGG 211  
 QY 21 GluIleAlaGlnValaAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
 |||||  
 DB 212 GAAATGCCCCAGGTGGACCCCAAGAACCATTCAGATGGATCTTCCGGATCAATCCA 271  
 QY 41 AspGlySerGlnSerValaIglValaProtyrAlaArgSerGluAlaHisLeuThrGlu 60  
 |||||  
 DB 272 GATGGAGCCAGCATCTGCTGGAGGTGCTTATGCCCCCTCAGAGGCCACCTCACAGAG 331  
 QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGlnGlnIleAspProSerThr 80  
 |||||  
 DB 332 CTGCTGAGAGCATATGTGACCGCATGAAGAGATGGGGAACAGATTATCTTCACCC 391  
 QY 81 HisArgLysAsnTyrValaArgValaIglValaGlnGlyGlnSerGluLeuAspLeu 100  
 |||||  
 DB 392 CATCGCAAGAACTACTAGCTAGTGGGCCGGAATGGAAATCCAGTACAGTGCCTA 451  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlnSerIle 120  
 |||||  
 DB 452 CAAGGCATCCGATGACACATATTAACGCCACCTCAAGTTTGGTGTAGAGCATTT 511  
 QY 121 ValGluGluTyrGluAspGluLeuIleGlnPheSerArgGluAlaAspAsnValLys 140  
 |||||  
 DB 512 GTGGAGGAATACAGAGATCACTGAATTCCTTCCCGAGAGCGTGCACATTTTAA 571  
 QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 |||||  
 DB 572 GACAAACTTTCAGTAAAGCAACAGATCTTTGTGACATGCCCCCTGCACATATCGCATGAT 631  
 QY 161 GluLeu 162  
 |||||  
 DB 632 GAGCTA 637

## RESULT 6

BE733900 719 bp mRNA linear EST 15-SEP-2000  
 BE733900  
 LOCUS 601568370F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:3843209 5',  
 DEFINITION mRNA sequence.  
 BE733900  
 ACCESSION BE733900  
 VERSION BE733900.1 GI:10147892  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 719)  
 NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: c9apbs-r@mail.nih.gov  
 Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
 Plate: LLC537 row: k column: 18  
 High quality sequence stop: 710.  
 Location/Qualifiers  
 1..719  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3843209"  
 /clone\_1db="NIH\_MGC\_21"  
 /tissue\_type="choriocarcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: placenta; Vector: pCMV-SPORT6; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCAGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 190 a 164 c 214 g 151 t  
 ORIGIN

## Alignment Scores:

Pred. No.:	8.4e-103	Length:	719
Score:	848.00	Matches:	162
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-10-082-502-20 (1-162) x BE733900 (1-719)

QY 1 ArgArSergInSArValaIglValaProtyrAlaArgSerGluAlaHisLeuThrGlu 20  
 |||||  
 DB 108 CGGAGGAGCCAGATCTCCACTGTGGAGCATGCCAGGCTGTGGATGAACTAGATGG 167  
 QY 21 GluIleAlaGlnValaAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
 |||||  
 DB 168 GAAATGCCCCAGGTGGACCCCAAGAACCATTCAGATGGATCTTCCGGATCAATCCA 227  
 QY 41 AspGlySerGlnSerValaIglValaProtyrAlaArgSerGluAlaHisLeuThrGlu 60  
 |||||  
 DB 228 GATGGAGCCAGCATCTGCTGGAGGTGCTTATGCCCCCTCAGAGGCCACCTCACAGAG 287  
 QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGlnGlnIleAspProSerThr 80  
 |||||  
 DB 288 CTGCTGAGAGCATATGTGACCGCATGAAGAGATGGGGAACAGATTATCTTCACCC 347  
 QY 81 HisArgLysAsnTyrValaArgValaIglValaGlnGlyGlnSerGluLeuAspLeu 100  
 |||||  
 DB 348 CATCGCAAGAACTACTAGCTAGTGGGCCGGAATGGAAATCCAGTACAGTGCCTA 407  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlnSerIle 120  
 |||||  
 DB 408 CAAGGCATCCGATGACACATATTAACGCCACCTCAAGTTTGGTGTAGAGCATTT 467  
 QY 121 ValGluGluTyrGluAspGluLeuIleGlnPheSerArgGluAlaAspAsnValLys 140  
 |||||  
 DB 468 GTGGAGGAATACAGAGATCACTGAATTCCTTCCCGAGAGCGTGCACATTTTAA 527  
 QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 |||||  
 DB 528 GACAAACTTTCAGTAAAGCAACAGATCTTTGTGACATGCCCCCTGCACATATCGCATGAT 587  
 QY 161 GluLeu 162  
 |||||  
 DB 588 GAGCTA 593

## RESULT 7

LOCUS	BG824549	742 bp	mRNA	linear	EST 22-MAY-2001
DEFINITION	603722845491 NIH_MGC_15 Homo sapiens CDNA clone IMAGE:4868100 5',				
ACCESSION	mRNA sequence.				
VERSION	BG824549				
KEYWORDS	BG824549.1	GI:14172136			
SOURCE	EST.				
ORGANISM	human.				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 742)				
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a> Tissue Procurement: ATCC CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) DNA Sequencing by: NIH Intramural Sequencing Center Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: L1CML737 row: k column: 13 High quality sequence stop: 740.				
FEATURES	Location/Qualifiers				
source	1..742				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:4868100"				
	/clone_lib="NIH_MGC_15"				
	/tissue_type="adenocarcinoma cell line"				
	/lab_host="DH10B (phage-resistant)"				
	/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adapter: GGGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"				
BASE COUNT	180 a	184 c	233 g	144 t	1 others
ORIGIN					
Alignment Scores:					
Pred. No.:	8,81e-103	Length:	742		
Score:	848.00	Matches:	162		
Percent Similarity:	100.00%	Conservative:	0		
Best local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	12	Gaps:	0		
US-10-082-502-20 (1-162) x BG824549 (1-742)					
Oy	1	ArgAGSerGlnAAspLeuHhIScYsgLIaCSArGAlaLeuValAspGluLengLutrp	20		
Db	201	CGGAGGACCCAGGATCTCCACCTGTCGACCATGACGCGCTGTGTGATGAACATAGAATGG	260		
Oy	21	GlutIleAGlnValAspProLysLysThrIleGlnMetGlySerPhaArgIleAspPro	40		
Db	261	GAATTTGCCAGAGTGGACCCCAAGAAAGCCATTCAGATGGGATCTTTCGGATCAATCCA	320		
Oy	41	AspGlySerGlnSerValValGluValProTYrAlaArgSerGluAlaHisLeuThrGlu	60		
Db	321	GATGCGACACGATCATGTGTGTGAGAGTGGCTTATATGCCCTCAGAGGCCACCTCACAAG	380		
Oy	61	LeuLeuGluGluIleCysAspArgMetLysGluTYrGlyGluGlnIleAspProSerThr	80		
Db	381	CTGCGGAGGAGATATGTGACCGGATGCAAGAGATGTGGGAACAGATTGATCTTCACCC	440		
Oy	81	HisArgLysAsnTYrValArgValValGluTYrArgAsnGlyLysSerGlnLeuAspLeu	100		

DB	441	CATCCAGAACTACGTACGTCGTAGTGGGCCCGAATTGAGANTGCAGATGCAGTAACGTGACGCA	500
OY	101	Gtnglyllleatyltleaspserrspileserrglythrleuleysphalascyglusertlle	120
DB	501	GACAACTTGACGTAAAGCAGACAGATCTTTGTGACACATGCCCTGCACATATCGCATGT	680
OY	121	VAlgtugltvrglAspdluleullegiuphepneserArgGluaLAspasnValys	140
DB	561	GTGGAGGATACGACAGATGACTCATTTGANTCTTTTTCCCGAGAGCTTGACATGTATAA	620
OY	141	AsplylsLeucysSerLyArqtThraspleucysAsprhlalaLeuhnlIseserhasp	160
DB	621	GACAACCTTGACGTAAAGCAGACAGATCTTTGTGACACATGCCCTGCACATATCGCATGT	680
OY	161	Gluten 162	
DB	681	GAGCTA 686	
RESULT 8			
LOCUS	BE791763	745 bp	mRNA linear EST 20-SEP-2000
DEFINITION	BE791763	60158158661 NIH_MGC_7 Homo sapiens CDNA clone IMAGE:3936141 5'	
ACCESSION	BE791763	mRNA sequence.	
VERSION	BE791763.1	GI:10212961	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.ncl.nih.gov/		
JOURNAL	1 (bases 1 to 745)		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgaabs-remail.nih.gov		
	Tissue procurement: DCD/DMP		
	CDNA library preparation: Ling Hong/Rudin Laboratory		
	CDNA library arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov		
	Plate: LINC79 row: k column: 22		
	High quality sequence stop: 739.		
FEATURES			
source	Location/Qualifiers		
	1..745		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3936141"		
	/clone_lib="NIH_MGC_7"		
	/tissue_type="small cell carcinoma"		
	/cell_line="MGC3"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: lung; Vector: pORF7; Site: 1: XhoI; Site_2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."*		
BASE COUNT	182 a 180 c 237 g 146 t		
ALIGNMENT SCORES:			
Pred. No.: 8	866-103	Length: 745	
Score: 848.00		Matches: 162	
Percent Similarity: 100.00%		Conservative: 0	
Best local Similarity: 100.00%		Mismatches: 0	
Query Match: 100.00%		Indels: 0	
DB: 12		Gaps: 0	

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 DB 198 CGAGGAGCCGAGATCTCCACTGTGAGCATGCGAGGCTGTGGTGAATGATGATG 257  
 QY 21 GlnIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
 DB 258 GAATTCGCCAGGTGGACCCCAAGAGACCATTCAGATGGATCTTCCGATCAATCCA 317  
 QY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
 DB 318 GATGGCAGCCAGTCAGTGTGGAGGTCCCTTATGCCCGCTCAAGGCCCACTCAGAG 377  
 QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 DB 378 CTCTGGAGGAGATATGTGACCGATGAGAGATATGGGAGACAGATTGATCTTCAC 437  
 QY 81 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyGluSerSerGluLeuAspLeu 100  
 DB 438 CATCGCAAGAACTACGTAGCTGTAGTGGCCGAGATGAGAAATCCAGTGAAGTGA 497  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 DB 498 CAAGGATCCGAAATCGATCAGATATTAGCGGACCCCTCAAGTTGCGTGAAGCAT 557  
 QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
 DB 558 GTGGAGATACGAGATGATACATCTTGAATCTTTCCGAGAGGCTACATGTTAA 617  
 QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 DB 618 GACAAACTTTCAGATGAGAGACAGATCTTTGTGACCATGCCCTGCACATATCGATGAT 677  
 QY 161 GluLeu 162  
 DB 678 GAGCTA 683

RESULT 9  
 LOCUS BQ437660 930 bp mRNA linear EST 24-MAY-2002  
 DEFINITION AGENCOURT\_7897085 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6159838  
 ACCESSION BQ437660  
 VERSION BQ437660.1 GI:21176736  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 930)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: ATCC/DCIT/DMP  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LMNL at:  
 http://image.llnl.gov  
 Plate: LLM13509 row: 1 column: 23  
 High quality sequence stop: 515.  
 Location/Qualifiers  
 1. .930  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6159838"  
 /clone\_id="NIH\_MGC\_72"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: skin; Vector: pCMV-SPORT6; Site: 1; NCI;"

FEATURES  
 source

BASE COUNT 244 a 221 c 257 g 208 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,23e-102 Length: 930  
 Score: 848.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 14 Gaps: 0  
 US-10-082-502-20 (1-162) x BQ437660 (1-930)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 DB 10 CGAGGAGCCGAGATCTCCACTGTGAGCATGCGAGGCTGTGGTGAATGATGATG 69  
 QY 21 GlnIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
 DB 70 GAATTCGCCAGGTGGACCCCAAGAGACCATTCAGATGGATCTTCCGATCAATCCA 129  
 QY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
 DB 130 GATGGCAGCCAGTCAGTGTGGAGGTCCCTTATGCCCGCTCAAGGCCCACTCAGAG 189  
 QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 DB 190 CTCTGGAGGAGATATGTGACCGATGAGAGATATGGGAGACAGATTGATCTTCAC 249  
 QY 81 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyGluSerSerGluLeuAspLeu 100  
 DB 250 CATCGCAAGAACTACGTAGCTGTAGTGGCCGAGATGAGAAATCCAGTGAAGTGA 309  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 DB 310 CAAGGATCCGAAATCGATCAGATATTAGCGGACCCCTCAAGTTGCGTGAAGCAT 369  
 QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
 DB 370 GTGGAGATACGAGATGAGATCATCTTGAATCTTTCCGAGAGGCTACATGTTAA 429  
 QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 DB 430 GACAAACTTTCAGATGAGAGACAGATCTTTGTGACCATGCCCTGCACATATCGATGAT 489  
 QY 161 GluLeu 162  
 DB 490 GAGCTA 495

RESULT 10  
 LOCUS BM552907 1117 bp mRNA linear EST 20-FEB-2002  
 DEFINITION AGENCOURT\_5572552 NIH\_MGC\_41 Homo sapiens cDNA clone IMAGE:5466942  
 ACCESSION BM552907  
 VERSION BM552907.1 GI:18791172  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 1117)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: DCTD/DMP  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L16M1968 row: c column: 07  
High quality sequence stop: 709.

## FEATURES

Location/Qualifiers

1. 1117  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5466942"  
/clone.lib="NIH.MGC.41"  
/tissue\_type="melanotic melanoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using Zap-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH-MGC Library."

BASE COUNT 299 a 258 c 320 g 239 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.: 1.62e-102 Length: 1117  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-10-082-502-20 (1-162) x BM552907 (1-1117)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTTP 20  
DB 208 CGGAGAGCCAGCATCTCCACCTGTGAGCATGAGGCGCTCTGTGATGATGACATGAGATGG 267  
QY 21 GGUUUAAGUValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
DB 268 GAAATGCCCCAGGTGAGACCCCAAGAACACATTCAGATGGGATCTTCCGATCAATCCA 327  
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnValHisLeuThrGlu 60  
DB 328 GATGGAGCCAGTCACTGTGTGAGGTGCTTATGCGCTCAAGAGCCACCTCACAGAG 387  
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
DB 388 CTCCTGAGAGATATGTATCCGATGAAAGAGATATGGGAAACAGATTGATCTTCCACC 447  
QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerSerGluLeuAspLeu 100  
DB 448 CATCGCAAGAACTACCTAGCTAGTGTAGTGGCCGCAATGAGAAATCCAGTGAATCGACCTA 507  
QY 101 GlnGlyIleArgIleAspSerAspLysSerGlyThrLeuLysPheValAspGluSerIle 120  
DB 508 CAAGGATCCGATTCAGTCACTGTATTAAGCGACCTCAAGTTTCGCTGTGAGACATT 567  
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluValAspAsnValLys 140  
DB 568 GTGGAGAGATACGAGATTAATCTGTAATCTTTTCCCGAGAGGCTGACATGTTAAA 627  
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
DB 628 GACAAACCTTTCAGATTAAGGAGACAGATCTTTGTGACCATGACCTGCACATATCGCATGAT 687  
QY 161 GluLeu 162  
DB 688 GAGCTA 693

RESULT 11  
BI791489/c

LOCUS BI791489 599 bp mRNA linear EST 11-MAR-2002  
DEFINITION Id98f09.x1 HR85 Islet Homo sapiens cDNA clone IMAGE:5085832 3'  
sequence.

## ACCESSION

BI791489

## VERSION

BI791489.1 GI:15819214

## KEYWORDS

EST.

## SOURCE

human.

## ORGANISM

Homo sapiens

## REFERENCE

Mammalia: Eutheria: Primates; Catarrhini; Hominoidea: Homo.

## AUTHORS

1 (bases 1 to 599)

## TITLE

Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,  
Lewisha, I., Scaevola, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistein, A.,  
Schmitt, A., Thelings, B., Rutter, E., Ronko, I., Bennett, J., Cardenas  
, M., Gibbons, M., McCann, R., Cole, R., Tsagarashvili, R., Williams, T.,  
Jackson, Y. and Bowers, Y.

## JOURNAL

Endocrine Pancreas Consortium  
Unpublished (2000)

## COMMENT

Other\_ESTS: Id98f09.y1

## CONTACT

Contact: Douglas Melton, Klaus H. Kaestner, &amp; Hiroshi Inoue

## DEPT

Harvard University, Howard Hughes Medical Institute

## ADDRESS

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138

## TELEPHONE

Tel: 617-495-1812

## FAX

Fax: 617-495-8557

## EMAIL

Email: dmelton@lohp.harvard.edu

## LIBRARY

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@wustl.edu)  
Seq primer: -40UP from G1bco

## HIGH QUALITY

High quality sequence stop: 399.

## LOCATION

Location/Qualifiers

## ORIGIN

1. 599

## ALIGNMENT

/organism="Homo sapiens"

## SCORE

/db\_xref="taxon:9606"

## PERCENT

/clone="IMAGE:5085832"

## BEST LOCAL

/clone.lib="HR85 Islet"

## QUERY MATCH

/tissue\_type="Purified pancreatic islet"

## BASE COUNT

/lab\_host="DH10B"

## ALIGNMENT

/note="Organ: Pancreas; Vector: pBluescript SK(-); Site:1:  
NotI; Site:2: XhoI; cDNA made by oligo-dT priming.  
Size-selected on agarose gel. Average insert size -1kb. 5'  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Permutt Lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."

## PRED. NO.

2.99e-102

## SCORE

843.00

## PERCENT

100.00%

## BEST LOCAL

100.00%

## QUERY MATCH

98.41%

## DB:

13 Gaps: 0

US-10-082-502-20 (1-162) x BI791489 (1-599)

QY 2 ArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTTPG 21  
DB 599 AGAGCCAGAGATCTCCACCTGTGAGCATGAGGCGCTCTGTGATGATGACATGAGATGG 540  
QY 22 IleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnProAsp 41  
DB 539 ATTGCCAGGTGAGACCCCAAGAACACATTCAGATGGGATCTTTCGAGATCAATCAGAT 480

QY 42 GlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGluLeu 61  
 |||  
 Db 479 GGCAGCCAGCTAGTGTGGAGGTCTTATGCCCGCCAGAGGCCACCTCACAGACCTG 420  
 QY 62 LeuGluGluIleCysAspArgMetGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 81  
 |||  
 Db 419 CTGGAGGAGATGATGTGACCGGATGAAGAGATATGGGAGAACAGATTGATCTTCACCCAT 360  
 QY 82 ArgGlySerGlnValAlaGluValGlyArgAsnGlyGluSerSerGluLeuAspLeuGln 101  
 |||  
 Db 359 CGCAGAACACTGACCTGATGAGTGGCCGGAATGAGAACTCAGTGAATCGAGACCTACAA 300  
 QY 102 GlyIleArgIleAspSerAspIleSerGlyThrLeuIleAspGlyGlySerIleVal 121  
 |||  
 Db 299 GGCATCGCATTCACCTCAGATATGAGCGACCCCTCAAGTTGCGTGTGAGACATTTGTG 240  
 QY 122 GluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLysAsp 141  
 |||  
 Db 239 GAGGAATACGAGATGAACTGATTAATCTTCCGAGAGGCTGACATGTTAAAGAC 180  
 QY 142 LysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAspGlu 161  
 |||  
 Db 179 AATCTTGAGTAAGCAGACAGATCTTTGTGACCATGCCCTGCATATCGCATGTAG 120  
 QY 162 Leu 162  
 |||  
 Db 119 CTA 117  
 RESULT 12  
 BQ421827 880 bp mRNA linear EST 23-MAY-2002  
 LOCUS DEFINITION AGENCOURT\_7801623 NIH\_MGC\_72 Homo sapiens CDNA clone IMAGE:6050568  
 BQ421827 5', mRNA sequence.  
 BQ421827.1 GI:21117142  
 EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 880)  
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: ATCC/DCTP/DRP  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLMJ3303 row: a column: 01  
 High quality sequence stop: 580.  
 Location/Qualifiers  
 1..880  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6050568"  
 /clone\_1lb="NIH\_MGC\_72"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: skin; Vector: pCMV-Sport6; Site: 1: Not;  
 Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 2 kb. Library constructed by Life  
 Technologies."  
 BASE COUNT 238 a 208 c 256 g 176 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.2e-102 Length: 880  
 Score: 842.00 Matches: 161

Percent Similarity: 99.388  
 Best Local Similarity: 99.388  
 Query Match: 99.29%  
 Db: 14  
 Gaps: 0  
 US-10-082-502-20 (1-162) x BQ421827 (1-880)  
 QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 |||  
 Db 162 CGAGAGAGCAGAGACTCCACGTGTGAGCATGCAGGCGCTGTGGATGACATGAAATGG 221  
 QY 21 GluIleAlaGluValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 40  
 |||  
 Db 222 GAATATGCCAGAGTGGAGCCCAAGAAAGACATTCAGATGGGATCTTCCGATCAATCCA 281  
 QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 |||  
 Db 282 GATGGCAGCCAGTACGTGTGTGAGGTGCTTATGCCGCTCAGAGGCCACCTCAGAG 341  
 QY 61 LeuLeuGluGluIleCysAspArgMetGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 80  
 |||  
 Db 342 CTGCTGAGAGAGATATGTGACCGGATGAGAGATGGGGAACAGATGATGATCTCCACC 401  
 QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGluSerSerGluLeuAspLeu 100  
 |||  
 Db 402 CATGCCAAGAACTACGTACGTGTAGTGGCCGTAAATGAGAACTCAGTGAATGAGCTTA 461  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 |||  
 Db 462 CAAGGCATCCCAATCGACTGATATATTAGCGGACCCCTCAAGTTGGTGTGAGAGCAT 521  
 QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
 |||  
 Db 522 GTGGAGGAATACGAGAGATGAACTGATTAATCTTCCGAGAGGCTGACAAATGTANA 581  
 QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 |||  
 Db 582 GACAAACTTTCAGTAAGCAGAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 641  
 QY 161 GluLeu 162  
 |||  
 Db 642 GAGCTA 647  
 RESULT 13  
 B1550436 783 bp mRNA linear EST 05-SEP-2001  
 LOCUS DEFINITION 603192905F1 NIH\_MGC\_95 Homo sapiens CDNA clone IMAGE:5264328 5',  
 B1550436 mRNA sequence.  
 B1550436  
 VERSION B1550436.1 GI:15437748  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 783)  
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki  
 Toshiyuki and Piero Carninci (RIKEN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLMJ1666 row: 1 column: 01  
 High quality sequence stop: 730.  
 Location/Qualifiers  
 1..783  
 FEATURES  
 source

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5264328"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.5 kb and
normalized to 80% 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      209 a      189 c      225 g      160 t
ORIGIN

Alignment Scores:
Pred. No.:      1,12e-101      Length:      783
Score:          840.00      Matches:      161
Percent Similarity: 99.38%      Conservative: 0
Best Local Similarity: 99.38%      Mismatches: 1
Query Match:     99.06%      Indels:      0
DB:              13      Gaps:      0

US-10-082-502-20 (1-162) x B1550436 (1-783)
OY      1      ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
        |||||
DB      144      CGGAGAGAGCCAGAGATCCACCTGTGAGCATGCGAGGCTCTGTGATGATGACATGAGATGG 203
OY      21      GluIleAlaGlnValAspProLysLysThrIleGlnMetClySerPheArgIleAsnPro 40
        |||||
DB      204      GAAATGGCCAGGTGGAGCCCAAGAACCATTCAGATGGGATCTTCGGATCAATCA 263
OY      41      AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGlnLysIleLeuThrGlu 60
        |||||
DB      264      GATGGAGCCAGCATGATGAGTGAGGCTGTATGCCGTAAAGAGCCCACTCAGACAGAG 323
OY      61      LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAsnProSerThr 80
        |||||
DB      324      CTGCTGGAGAGATATGTGACCGGATGAGCATGAGGAGACAGATGATCTTCACAC 383
OY      81      HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 100
        |||||
DB      384      CATCGCAAACTACCTAGCTAGTGTAGTGCGGCGAATGAGAAATCCAGTGAATCGACCTA 443
OY      101      GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
        |||||
DB      444      CAAGGCATCCGAATGACATGATATATAGCGGACACCTCAAGTTTCCGTGTGAGACCAT 503
OY      121      ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluLysAspAsnValLys 140
        |||||
DB      504      GTGGAGATACGAGATGAATCAATCTTATTTCTTTCCCGAGAGGCTACATGTTTANA 563
OY      141      AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
        |||||
DB      564      GACAAACTTTGCAATGAGGAAACAGATCTTTGTGACCATGCGCCTGCACATATGATGAT 623
OY      161      GluLeu 162
        |||||
DB      624      GAGCTA 629

RESULT 14
LOCUS      B0421069      868 bp      mRNA      linear      EST 23-MAY-2002
DEFINITION AGERCOURT.7911956 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6009949
5', mRNA sequence.
ACCESSION      B0421069
VERSION      B0421069.1 GI:21116384
KEYWORDS
SOURCE      human.
```

```
ORGANISM      Homo sapiens
REFERENCE      Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
TITLE      1 (bases 1 to 868)
JOURNAL      NIH-MGC http://img.nci.nih.gov/.
COMMENT      National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-riemail.nih.gov
Tissue Procurement: DCTD/DRP/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM13197 row: d column: 14
High quality sequence stop: 542.
FEATURES
source
location/Qualifiers
1..868
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6009949"
/clone_lib="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies."
BASE COUNT      232 a      203 c      257 g      168 t      8 others
ORIGIN

Alignment Scores:
Pred. No.:      6.09e-101      Length:      868
Score:          835.00      Matches:      160
Percent Similarity: 98.77%      Conservative: 0
Best Local Similarity: 98.77%      Mismatches: 2
Query Match:     98.47%      Indels:      0
DB:              14      Gaps:      0

US-10-082-502-20 (1-162) x B0421069 (1-868)
OY      1      ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
        |||||
DB      190      CGGAGAGAGCCAGAGATCCACCTGTGAGCATGCGAGGCTCTGTGATGATGACATGAGATGG 249
OY      21      GluIleAlaGlnValAspProLysLysThrIleGlnMetClySerPheArgIleAsnPro 40
        |||||
DB      250      GAAATGGCCAGGTGGAGCCCAAGAACCATTCAGATGGGATCTTCGGATCAATCA 309
OY      41      AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGlnLysIleLeuThrGlu 60
        |||||
DB      310      GATGGAGCCAGCATGATGAGTGAGGCTGTATGCCGTCAAGAGCCCACTCAGACAGAG 369
OY      61      LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAsnProSerThr 80
        |||||
DB      370      CTGCTGGAGAGATATGTGACCGGATGAGCATGAGGAGACAGATGATCTTCACAC 429
OY      81      HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 100
        |||||
DB      430      CATCGCAAACTACCTAGCTAGTGTAGTGCGGCGAATGAGAAATCCAGTGAATCGACCTA 489
OY      101      GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
        |||||
DB      490      CAAGGCATCCGAATGACATGATATATAGCGGACACCTCAAGTTTCCGTGTGAGACCAT 549
OY      121      ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluLysAspAsnValLys 140
        |||||
DB      550      GTGGAGATACGAGATGAATCAATCTTATTTCTTTCCCGAGAGGCTACATGTTTANA 609
OY      141      AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
        |||||
```



DB 610 GACAACTTTGACGACGACGACATCTTTGTGACCATGCCCTGCACATATGCATGAT 669  
QY 161 GIUleu 162  
DB 670 GAGCTA 675  
RESULT 15  
BG704443  
LOCUS  
DEFINITION BG704443 791 bp mRNA linear EST 07-MAY-2001  
602686633F1 NIH\_MGC\_95 Homo sapiens CDNA clone IMAGE:4821008 5',  
mRNA sequence.  
ACCESSION BG704443  
VERSION BG704443  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 791)  
AUTHORS NIH-MGC http://mgs.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@strs-remail.nih.gov  
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL0727 row: a column: 09  
High quality sequence stop: 785.  
FEATURES  
Source  
Location/Qualifiers  
1..791  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4821008"  
/clone\_1lb="NIH\_MGC\_95"  
/tissue\_type="hippocampus"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pBluescript (modified  
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtagag  
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',  
size-selected for average insert size 2.5 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."  
BASE COUNT 187 a 196 c 256 g 152 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 9.83e-101 Length: 791  
Score: 833.00 Matches: 160  
Percent Similarity: 99.38% Conservative: 1  
Best Local Similarity: 98.77% Mismatches: 1  
Query Match: 98.23% Indels: 0  
Gaps: 0  
US-10-082-502-20 (1-162) x BG704443 (1-791)  
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
DB 259 CGGAGGAGCCAGAGATCTCCAGATGTCGACAGCAGCGCTCTGGTGATGAAGTAATGG 318  
QY 21 GluLeaGlnValAspProGlySerThrIleGlnMetGlySerPheArgIleAsnPro 40  
DB 319 GAATATGCCAGGAGCCCAAGAGACCATTCAGATGGGATCTTCCGATCAATCCA 378  
QY 41 AspGlySerGlnSerValIleGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60

DB 379 GATGGCAGCCAGTCAGTGGTGAGGTGCCCTTATGCCCGCTCAGAGGCCACACCTCACAGAG 438  
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
DB 439 CTGCTGGAGGAGATATGTGACCGGATGMAAGAGTATGGGAAACAGATTGATCTTCACACC 498  
QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerSerGluLeuAspLeu 100  
DB 499 CATCGCAAGAACTACGTACGTAGTGCGCCGGAATGAGAAATCCAGTCACTGGAGCCTA 558  
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 559 CAAGGATCCAAATCCACTCAGTATTACGGCACCCCTCAAGTTGGCTGTGAGACAT 618  
QY 121 ValGluGluTyrGluAspLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
DB 619 GTGGAGGAATACGAGGATGACATGTAATCTTTCCGAGAGGCTGCACAACTTTAA 678  
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
DB 679 GACAACTTTGGAGTAAAGCAACAGATCTTTGTGACCATGCCCTGCACATATGCATGAT 738  
QY 161 GIUleu 162  
DB 739 GAGCTA 744

Search completed: March 31, 2003, 04:51:11  
JDB time : 1041.65 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 07:02:46 ; Search time 48.9767 Seconds

(without alignments)  
2814.757 Million cell updates/sec

Title: US-10-082-502-20

Perfect score: 848  
Sequence: 1 RNSQDLHCACGACRALVDELM.....LCSKRRDLDHALHSHDEL 162

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=xip  
-O/cg2\_1/USPTO.spool/US10082502/runat\_24032003\_135103\_6870/app.query.fasta.1.1308  
-DB=Published.Applications.NA -OFMT=fastap -SUFFIX=rrmpd -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdd -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAILEN=2000000000 -USER=US10082502 -CGN\_1.1.145 -runat\_24032003\_135103\_6870  
-NCPU=6 -ICPU=3 -NO\_XLPHY -NO\_MMAP -LARGOQUERY -NEG\_SCORES=0 -WAIT -LONGIOS  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THEADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published.Applications\_NA:\*

1: /cg2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cg2\_6/ptodata/2/pubpna/PCR\_NEM\_PUB.seq:\*  
3: /cg2\_6/ptodata/2/pubpna/US06\_NEM\_PUB.seq:\*  
4: /cg2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cg2\_6/ptodata/2/pubpna/US07\_NEM\_PUB.seq:\*  
6: /cg2\_6/ptodata/2/pubpna/PCITUS\_PUBCOMB.seq:\*  
7: /cg2\_6/ptodata/2/pubpna/US08\_NEM\_PUB.seq:\*  
8: /cg2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cg2\_6/ptodata/2/pubpna/US09\_NEM\_PUB.seq:\*  
10: /cg2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cg2\_6/ptodata/2/pubpna/US10\_NEM\_PUB.seq:\*  
12: /cg2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cg2\_6/ptodata/2/pubpna/US60\_NEM\_PUB.seq:\*  
14: /cg2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	848	100.0	814	10 US-09-864-864-309
2	848	100.0	1210	9 US-10-028-072-173
3	848	100.0	1210	9 US-10-121-049-173
4	848	100.0	1210	9 US-10-123-904-173

5	848	100.0	1210	9 US-10-140-470-173	Sequence 173, App
6	848	100.0	1210	9 US-10-175-746-173	Sequence 173, App
7	848	100.0	1210	9 US-10-176-918-173	Sequence 173, App
8	848	100.0	1210	9 US-10-176-921-173	Sequence 173, App
9	848	100.0	1210	9 US-10-137-865-173	Sequence 173, App
10	848	100.0	1210	9 US-10-140-474-173	Sequence 173, App
11	848	100.0	1210	9 US-10-142-311-173	Sequence 173, App
12	848	100.0	1210	9 US-10-143-114-173	Sequence 173, App
13	848	100.0	1210	9 US-10-140-002-173	Sequence 173, App
14	848	100.0	1210	9 US-10-142-419-173	Sequence 173, App
15	848	100.0	1210	9 US-10-123-262-173	Sequence 173, App
16	848	100.0	1210	9 US-10-142-423-173	Sequence 173, App
17	848	100.0	1210	9 US-10-121-050-173	Sequence 173, App
18	848	100.0	1210	9 US-10-141-755-173	Sequence 173, App
19	829	97.8	657	9 US-09-832-846-198	Sequence 398, App
20	154	18.2	564	10 US-09-864-864-228	Sequence 228, App
21	129	15.2	1457	10 US-09-846-573B-10	Sequence 10, App1
22	129	15.2	1470	9 US-10-036-041-58	Sequence 58, App1
23	129	15.2	1470	9 US-10-035-855-58	Sequence 58, App1
24	129	15.2	1470	9 US-10-174-590-479	Sequence 479, App
25	129	15.2	1470	9 US-10-176-758-479	Sequence 479, App
26	129	15.2	1470	9 US-10-175-737-479	Sequence 479, App
27	129	15.2	1470	9 US-10-173-706-479	Sequence 479, App
28	129	15.2	1470	9 US-10-175-738-479	Sequence 479, App
29	129	15.2	1470	9 US-10-175-752-479	Sequence 479, App
30	129	15.2	1470	9 US-10-176-482-479	Sequence 479, App
31	129	15.2	1470	9 US-10-176-757-479	Sequence 479, App
32	129	15.2	1470	9 US-10-176-913-479	Sequence 479, App
33	129	15.2	1470	9 US-10-180-552-479	Sequence 479, App
34	129	15.2	1470	9 US-10-180-557-479	Sequence 479, App
35	129	15.2	1470	9 US-09-931-836-58	Sequence 58, App1
36	129	15.2	1470	9 US-10-173-700-479	Sequence 479, App
37	129	15.2	1470	9 US-10-174-572-479	Sequence 479, App
38	129	15.2	1470	9 US-10-174-579-479	Sequence 479, App
39	129	15.2	1470	9 US-10-174-882-479	Sequence 479, App
40	129	15.2	1470	9 US-10-174-888-479	Sequence 479, App
41	129	15.2	1470	9 US-10-175-739-479	Sequence 479, App
42	129	15.2	1470	9 US-10-175-740-479	Sequence 479, App
43	129	15.2	1470	9 US-10-175-743-479	Sequence 479, App
44	129	15.2	1470	9 US-10-176-488-479	Sequence 479, App
45	129	15.2	1470	9 US-10-176-492-479	Sequence 479, App

#### ALIGNMENTS

RESULT 1  
US-09-864-864-309  
Sequence 309, Application US/09864864  
Patent No. US20020102679A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiansheng  
APPLICANT: Matcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Secrist, Heather  
APPLICANT: Lodes, Michael J.  
APPLICANT: Algate, Paul A.  
APPLICANT: Fling, Steve P.  
APPLICANT: Mannion, Jane  
APPLICANT: Benson, Darin R.  
APPLICANT: Carter, Patrick  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.523  
CURRENT APPLICATION NUMBER: US/09/864,864  
CURRENT FILING DATE: 2001-05-23  
NUMBER OF SEQ ID NOS: 341  
SOFTWARE: Corixa Invention Disclosure Database  
SEQ ID NO 309  
LENGTH: 814  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-864-864-309

Alignment Scores:  
 Pred. No.: 2,43e-102 Length: 814  
 Score: 848.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-082-502-20 (1-162) x US-09-864-864-309 (1-814)

OY 1 ARGARGSGLNAPLPHUHCYSGLYALACYSARGALALEUVALASPLUENGLUTP 20  
 DB 205 CGAGAGAGCCAGGATCTCCACTGTGAGCATGGCGCTCTGGTGAGTAAGATG 264  
 OY 21 GUILEALAGLVALAPPROLYSTHRIEGLIMETGLSERPHEATGLAENPRO 40  
 DB 265 GAAATGGCCAGGTGGACCCAGAGACCATTCACATGGGATCTTCCGATCATCCA 324  
 OY 41 ASBGLYSERGLNSERVALGLUVALPROTYRALARGSERGLUALAHISTEUTHGLU 60  
 DB 325 GAATGAGCCAGCATGAGTGAGTGCCCTTATGCCCCGCTCAGAGGCCACCTCACAGAG 384  
 OY 61 LEULEUGLUGLUECYSAPARGMETLYSGLUITYRGLUGLUGLHLEAPPROSERTHR 80  
 DB 385 CTCCTGAGAGCATATGTCACCGCATGAGGAGTATGGGAAACAGATTGCTCCACC 444  
 OY 81 HIASRGLYASNTYRVALARGVALVALGLYARGANGLYGLUSERSERGLUEASPLEU 100  
 DB 445 CATCGAAGAACTACGTACGTAGTGAGGGCCGAAATGGAGAAATCCAGTAACTGACCTA 504  
 OY 101 GLINGLYILEARGILEASPSERASPILESERGLYTHLEULYSPHEALACYSGLUSERLIE 120  
 DB 505 CAAGGCATCCGATTCACATGATTTACGGCACCTCAAGTTCCGTGTGAGACATT 564  
 OY 121 VALGLUGLUTYRGLUASPLUENGLUEGLUPHESERFARGLUALASPAENVALLYS 140  
 DB 565 GTGGAGAGATACAGATGACATCATTTCTTTTCCGAGAGGCTGACATGTTTAA 624  
 OY 141 ASPLYSLEUCYSERLYARGLTHASPLEUCYSAPSHIALALEUHLISERHISASP 160  
 DB 625 GACAACTTTCAGTAAAGCAAGAGATCTTTGTGACCATGCCCCCTGACATATGCAATGAT 684  
 OY 161 GLULEU 162  
 DB 685 GAGCTA 690

RESULT 2  
 US-10-028-072-173  
 ; Sequence 173, Application US/10028072  
 ; Publication No. US20030004311A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Beresini, Maureen  
 ; APPLICANT: Deforge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Sherwood, Steven  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Collin K  
 ; APPLICANT: Wood, William  
 ; APPLICANT: Zhang  
 ; TITLE OF INVENTION:  
 ; FILE REFERENCE:  
 ; CURRENT APPLICATION NUMBER: US/10/028,072  
 ; CURRENT FILING DATE: 2001-12-19

PRIOR APPLICATION NUMBER: 60/049911  
 PRIOR FILING DATE: 1997-06-18  
 PRIOR APPLICATION NUMBER: 60/056974  
 PRIOR FILING DATE: 1997-08-26  
 PRIOR APPLICATION NUMBER: 60/059113  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059115  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059117  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059122  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059184  
 PRIOR FILING DATE: 1997-09-17  
 PRIOR APPLICATION NUMBER: 60/059263  
 PRIOR FILING DATE: 1997-09-18  
 PRIOR APPLICATION NUMBER: 60/059352  
 PRIOR FILING DATE: 1997-09-19  
 PRIOR APPLICATION NUMBER: 60/059588  
 PRIOR FILING DATE: 1997-09-19  
 PRIOR APPLICATION NUMBER: 60/059836  
 PRIOR FILING DATE: 1997-09-24  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/062285  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/062287  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/062814  
 PRIOR FILING DATE: 1997-10-24  
 PRIOR APPLICATION NUMBER: 60/062816  
 PRIOR FILING DATE: 1997-10-24  
 PRIOR APPLICATION NUMBER: 60/063045  
 PRIOR FILING DATE: 1997-10-24  
 PRIOR APPLICATION NUMBER: 60/063082  
 PRIOR FILING DATE: 1997-10-31  
 PRIOR APPLICATION NUMBER: 60/063127  
 PRIOR FILING DATE: 1997-10-24  
 PRIOR APPLICATION NUMBER: 60/063327  
 PRIOR FILING DATE: 1997-10-27  
 PRIOR APPLICATION NUMBER: 60/063329  
 PRIOR FILING DATE: 1997-10-27  
 PRIOR APPLICATION NUMBER: 60/063550  
 PRIOR FILING DATE: 1997-10-28  
 PRIOR APPLICATION NUMBER: 60/063561  
 PRIOR FILING DATE: 1997-10-28  
 PRIOR APPLICATION NUMBER: 60/063704  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063733  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063735  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063738  
 PRIOR FILING DATE: 1997-10-29  
 PRIOR APPLICATION NUMBER: 60/063755  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/064248  
 PRIOR FILING DATE: 1997-11-03  
 PRIOR APPLICATION NUMBER: 60/064809  
 PRIOR FILING DATE: 1997-11-07  
 PRIOR APPLICATION NUMBER: 60/065186  
 PRIOR FILING DATE: 1997-11-12  
 PRIOR APPLICATION NUMBER: 60/065846  
 PRIOR FILING DATE: 1997-11-17  
 PRIOR APPLICATION NUMBER: 60/066364  
 PRIOR FILING DATE: 1997-11-21  
 PRIOR APPLICATION NUMBER: 60/066453  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/066511  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/066770  
 PRIOR FILING DATE: 1997-11-24  
 PRIOR APPLICATION NUMBER: 60/069212

PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069278  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069334  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069694  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 60/072320  
PRIOR FILING DATE: 1998-01-23  
PRIOR APPLICATION NUMBER: 60/073612  
PRIOR FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: 60/074086  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/074092  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-02-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081695  
PRIOR FILING DATE: 1998-04-14  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081818  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082999  
PRIOR FILING DATE: 1998-04-24  
PRIOR APPLICATION NUMBER: 60/083222  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085149  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/086414  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/086430  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088730  
PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088741  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090538  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07

## Alignment Scores:

Pred. No.: 4,2e-102 Length: 1210  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-20 (1-162) x US-10-028-072-173 (1-1210)

QY 1 ArgatSerGlnAspLeuHnScySglYalAcysAtgAlaLeuValAspGluLeuGluTrp 20  
DB 603 CGGAGGACCCAGCATCTCCACGTGTGAGCATGCGCTGCGATGACCTAGAAATGG 662  
QY 21 GluIleAlaGlnValAspProLysIleThrIleGlnMetGlySerPheArgIleAsnPro 40  
DB 663 GAAATTTGCCAGGTGGACCCCAAGAAAGACCATTCAGATGGAGTCTTCCGGATCAATCCA 722  
QY 41 AspGlySerGlnSerValValGluValProTyraIleArgSerGluAlaHnIleuThrGlu 60  
DB 723 GATGCGACGACGTAGTGTGTGAGGTGCTTATGCCCGCTCGAGGCCACCTCAGAG 782  
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 80  
DB 783 CTGCTGGAGGAGATATGTGACCGGATGAAGAGTGTGGGAACAGATTGATCTCCAC 842  
QY 81 HsArgLysAsnTyValArgValValGlyArgAsnGlyGluSerGluLeuAspLeu 100  
DB 843 CATCCAGAAACCTAGCTAGCTAGTGTAGTGGCCGAAATGGAATTCAGTAAGTGAAC 902  
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 903 CAAGCATCCGAAATGAGTACATGATTTAGCGGCACCTCAAGTTGGCGTGAAGCAT 962  
QY 121 ValGluGluTyrglyAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
DB 963 GTGGAGGAATAGAGAGATGAATCTCATTTCTTCCCGAGAGCTGCACAAATGTAAA 1022  
QY 141 AspLysIleCysSerLysArgThrAspLeuCysAspPheIleAlaLeuHnIleSerHsAsp 160  
DB 1023 GCAAACTTTGCAATAGCAAGAAACAGATCTTTGTGACCATGCCCTGCACATATGCGATAT 1082

OY 161 Glutelu 162  
DB 1083 GAGCTA 1088

RESULT 3  
US-10-121-049-173  
; Sequence 173, Application US/10121049  
; Publication No. US20030022239A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Geriltsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Collin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C17  
; CURRENT APPLICATION NUMBER: US/10/121,049  
; PRIORITY FILING DATE: 2002-04-12  
; NUMBER OF SEQ ID NOS: 550  
; Prior Application removed - See File Wrapper or Palm  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-121-049-173

Alignment Scores:  
Pred. No.: 4,2e-102 Length: 1210  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-20 (1-162) x US-10-121-049-173 (1-1210)

OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
DB 603 CGAGGAGCCAGGATCTCCACTGTGAGCATGCGAGGCTCTGTGGTGAACAGAAATGG 662

OY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetCysPheArgIleAsnPro 40  
DB 663 GAATATGCCAGATGTGACCCCAAGAACCATTCAGATGGGATCTTCCGATCAATCCA 722

OY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
DB 723 GATGGGAGCCAGATCTCCACTGTGAGCATGCGAGGCTCTGTGGTGAACAGAAATGG 782

OY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGlnIleAsnProSerThr 80  
DB 783 CTCTGTGAGGAGATATGTGACCCGATCAAGAGATATGGGAGACCATTCCTCCACC 842

OY 81 HisArgLysAsnTyrValAlaArgValAlaGlyArgAsnGlyLysSerGluLeuAspLeu 100  
DB 843 CATCGCAAGACTACAGCTGTAGTGGGCGGAATGGAAATCCAGTGAAGTGAAGCTTA 902

OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 903 CAAGGATCGAATCGATCGATATTTAGGGGACACCTCAAGTTTGGCTGTGAGAGCAT 962

OY 121 ValGluGluIleTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
DB 963 GTGGAGGAATACGAGATGATGATGAAATCTTTTCCGAGAGCGTCAATGTTAAA 1022

OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
DB 1023 GACAACTTTTGCAGTAAAGCAGACAGATCTTGTGACCATGCTCCGCAATATGCGATGAT 1082

OY 161 Glutelu 162  
DB 1083 GAGCTA 1088

RESULT 4  
US-10-123-904-173  
; Sequence 173, Application US/10123904  
; Publication No. US20030022328A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Geriltsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Collin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C54  
; CURRENT APPLICATION NUMBER: US/10/123,904  
; PRIORITY FILING DATE: 2002-04-16  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-123-904-173

Alignment Scores:  
Pred. No.: 4,2e-102 Length: 1210  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-20 (1-162) x US-10-123-904-173 (1-1210)

OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
DB 603 CGGAGAGCCAGGATCTCCACTGTGAGCATGCGAGGCTCTGTGGTGAACAGAAATGG 662

OY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetCysPheArgIleAsnPro 40  
DB 663 GAATATGCCAGGATGTGACCCCAAGAACCATTCAGATGGGATCTTCCGATCAATCCA 722

OY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
DB 723 GATGGGAGCCAGATCTCCACTGTGAGCATGCGAGGCTCTGTGGTGAACAGAAATGG 782

OY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGlnIleAsnProSerThr 80  
DB 783 CTCTGTGAGGAGATATGTGACCCGATCAAGAGATATGGGAGACCATTCCTCCACC 842

OY 81 HsArgLysAsnTyrValArgValAlaGlyArgAsnGlyGlnSerSerGlnLeuAspLeu 100  
| | | | |  
Db 843 CATCCCAAGAACTAGCTAGTGTAGTGGCGGAATGGAGATCCAGTAATGGAGCTTA 902  
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuPheAlaCysGlnSerIle 120  
| | | | |  
Db 903 CAGGCATCCGAATGACTAGATATTTAGCGGCACCTCAAGTTGGCTGTAGAGCAATT 962  
OY 121 ValGlnGlyTyrGlnAspGlnLeuIleGlnPheSerArgGlnAlaAspAsnValLys 140  
| | | | |  
Db 963 GTGGAGGAATACGAGAGTAAGTAATCTTTCCCGAGAGCTGACAAATGTAA 1022  
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisasp 160  
| | | | |  
Db 1023 GACAAACTTTGCAGTACGAGACAGATCTTTGTGACATGCCCTGCACATATCGCATGAT 1082  
OY 161 GlnLeu 162  
| | | | |  
Db 1083 GAGCTA 1088

## RESULT 5

US-10-140-470-173  
; Sequence 173, Application US/10140470  
; Publication No. US2003002231A1  
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3330R1C160  
CURRENT APPLICATION NUMBER: US/10/140,470  
CURRENT FILING DATE: 2002-05-06  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-140-470-173

## Alignment Scores:

Pred. No.: 4.2e-102 Length: 1210  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 9

US-10-082-502-20 (1-162) x US-10-140-470-173 (1-1210)

OY 1 ArgArgSerGlnAspLeuHisCysGlnAlaGlyArgAlaLeuValAspGlnLeuGlnTyr 20  
| | | | |  
Db 603 CGAGAGACCCAGATGTCCACTGTGAGCATGCGAGGCTGTGTGATGAATCAATAATGG 662  
OY 21 GlnIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
| | | | |  
Db 663 GAATTTGCCAGGTGAGACCCCAAGAAAGACCAATTCAGATGGGATCTTTCCGGAATCAATCA 722

OY 41 AspGlySerGlnSerValAlaGlnValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
| | | | |  
Db 723 GATGCGACGACCAATCAGTGTGTGAGGTGCTTATGCCCGCTCAGAGGCCACCTCACAGAG 782  
OY 61 LeuLeuGlnGlnIleCysAspArgMetLysGlyTyrGlyGlnGlnIleAspProSerThr 80  
| | | | |  
Db 783 CTGCTGGAGAGATATGTGACCGGATGAGAGATGGGGAAACAAATGATCTTCCACC 842  
OY 81 HsArgLysAsnTyrValArgValAlaGlyArgAsnGlyGlnSerSerGlnLeuAspLeu 100  
| | | | |  
Db 843 CATCCCAAGAACTAGCTAGTGTAGTGGCGGAATGGAGATCCAGTAATGGAGCTTA 902  
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuPheAlaCysGlnSerIle 120  
| | | | |  
Db 903 CAGGCATCCGAATGACTAGATATTTAGCGGCACCTCAAGTTGGCTGTAGAGCAATT 962  
OY 121 ValGlnGlyTyrGlnAspGlnLeuIleGlnPheSerArgGlnAlaAspAsnValLys 140  
| | | | |  
Db 963 GTGGAGGAATACGAGAGTAAGTAATCTTTCCCGAGAGCTGACAAATGTAA 1022  
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisasp 160  
| | | | |  
Db 1023 GACAAACTTTGCAGTACGAGACAGATCTTTGTGACATGCCCTGCACATATCGCATGAT 1082  
OY 161 GlnLeu 162  
| | | | |  
Db 1083 GAGCTA 1088

## RESULT 6

US-10-175-746-173  
; Sequence 173, Application US/10175746  
; Publication No. US20030027270A1  
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3330R1C353  
CURRENT APPLICATION NUMBER: US/10/175,746  
CURRENT FILING DATE: 2002-06-19  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-175-746-173

## Alignment Scores:

Pred. No.: 4.2e-102 Length: 1210  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 9

US-10-082-502-20 (1-162) x US-10-175-746-173 (1-1210)

OY	1	ATGATGSEGTGAAAPLLENINLACSGUVALACSYATGALALEUVALAARGIULEGUTTP	20
DB	603	CGAGAGACCCAGNATCTCCATGTGGAGCATGACGAGGGCTGTGGTGAATGAACATGATGG	662
OY	21	GIUILEAGLVALAARPROLYSLYSTHILEGIMETGLYSERPEARGIIAASPRO	40
DB	663	GAAATGTCCCGAGGTGGACCCCGAAGACACCTTCAGATGGGATCTTTCCGGATCAATGCCA	722
OY	41	AARGLYSERGINSERVALVALGULVALPROTYRALARSGERGIUALNILEAUPHGLU	60
DB	723	GATGGCACCCAGTACGTGTGGAGGTGCTTATTCGCCGCTCAGAGGCCACCTCACAGAG	782
OY	61	LEULEUGIULILECSYASPARGMETLYSGIUTPTIYGLUGIUNIIAASPROSETH	80
DB	783	CTGGTGGAGGAGATATGTGACCGGATGAGAGATATGGGGAACAGATGTATCTTCACCC	842
OY	81	HISARGLYSERANTPTVALARGVALVALGVALARGSNGLYGLISERGIULEAARPLEU	100
DB	843	CATGCCAAGAACTCGATACGTGTAGTGTGGCCGAGATGGAAATCCATGAACTGGACCTA	902
OY	101	GLNGIYILEARGIIAASPSERASPILSESERGLYTHREULYSPHEALACYSGLUSERILE	120
DB	903	CAGGCACTCCGAAATCGACTGAGATATTAGCGGACCTCAAGTTTGGCTGTGAGAGATT	962
OY	121	VALGUGIUTPTGILUASRGILEUILLIGIUPHESERARGIUALAASPAVALYLS	140
DB	963	GTGGAGGAATCGAGGAGAACTATTGAATCTTTTCCGAGGGGTGCAATGTGTA	1022
OY	141	ASPLYSLEACSERYARGTTHASPLEUCYASPHISALALEUHSILESERHISASP	160
DB	1023	GACAACTTTCAGATGAACGAGACAGATCTTTGTGCACATGCCCTGCACATATCGCATGAT	108
OY	161	GLULEU 162	
DB	1083		
		GAGCTA 1088	
RESULT 7			
		US-10-176-918-173	
		: Sequence 173, Application US/10176918	
		: Publication No. US20030027275A1	
		: GENERAL INFORMATION:	
		: APPLICANT: Baker, Kevin P.	
		: APPLICANT: Beresini, Maureen	
		: APPLICANT: DeForge, Laura	
		: APPLICANT: Desnoyers, Luc	
		: APPLICANT: Filvaroff, Ellen	
		: APPLICANT: Gao, Wei-Qiang	
		: APPLICANT: Gerritsen, Mary E.	
		: APPLICANT: Goddard, Audrey	
		: APPLICANT: Godowski, Paul J.	
		: APPLICANT: Gurney, Austin L.	
		: APPLICANT: Sherwood, Steven	
		: APPLICANT: Smith, Victoria	
		: APPLICANT: Stewart, Timothy A.	
		: APPLICANT: Tumas, Daniel	
		: APPLICANT: Watanabe, Colin K	
		: APPLICANT: Wood, William	
		: APPLICANT: Zhang, Zemin	
		: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC	
		: FILE REFERENCE: P3330R1C382	
		: CURRENT APPLICATION NUMBER: US/10/176,918	
		: CURRENT FILING DATE: 2002-06-20	
		: Prior Application removed - See file Wrapper or Palm	
		: NUMBER OF SEQ ID NOS: 550	
		: SEQ ID NO 173	
		: LENGTH: 1210	
		: TYPE: DNA	
		: ORGANISM: Homo Sapien	
		: US-10-176-918-173	
Alignment Scores:			
		4.2e-102	Length: 1210
		Pred. NO.:	

Score:	848.00	Matches:	152
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0
US-10-082-502-20 (1-162) x US-10-176-918-173 (1-1210)			
QY	1 ArgArgSerGlnAspIleuHHisCysGlyAlaCysArgAlaIleuValAspGluLeuGluTrp 20		
DB	603 CGGAGACCCAGGATCTCCATGTGGAGCATGCGGGCTGTGGATGAACTGAAATGG 662		
QY	21 GluIleAlaGlnValAspProTyrLysThrIleGlnMetGlySerPheArgIleAsnPro 40		
DB	663 GAAATGGCCAGGTGGAGACCCCAAGAGACCATTCAGATGGGATCTTCCGGATCAATCCA 722		
QY	41 AspGlySerGlnSerValValGlnValProTyrAlaIarGserGlnAlaHisIleuThrGlu 60		
DB	723 GATGGCACCCGCTAGTGGTGGAGCTCCCTTAATCCCGCTCAAGAGCCCACTCAGAG 782		
QY	61 LeuLeuGluGlnIleCysAspArgMetLysGluTyrGlyGlnIleAspProSerThr 80		
DB	783 CTGCTGGAGGAGATATGTGACCCGATCAAGAGATAGGGAGAACAGATATCTCTCCACC 842		
QY	81 HisArgLysAsnTyrValArgValValGlyAlaArgAsnGlyLysSerSerGluLeuAspLeu 100		
DB	843 CATCCCAAGAACTGACGTACGTGTGTGGCCGGATGTGAATGAAATCCAGTGAACCTGACCTA 902		
QY	101 GlnGlyIleAlaGlyIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlnSerIle 120		
DB	903 CAAGCATCCCAATCGATCGATCGATTATTAAGCGGCGACCTCAAGTTTGGCTGTGAGACATT 962		
QY	121 ValGluIuTyrGluAspGluLeuIleGluPheSerAlaGluAlaAspAsnValLys 140		
DB	963 GTGGAGAAATACGAGGATGAATCATTTGAATCTTTTCCCGAAGAGCTGCACATGTTAA 1022		
QY	141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaIleuHisIleSerHisAsp 160		
DB	1023 GACAAACTTTCAGTAAAGCGAAGACATCTTGTGACCATCTCCCTGCACATATGCATGAT 108		
QY	161 GluLeu 162		
DB	1083 GAGCTA 1088		
RESULT 8			
US-10-176-921-173			
Sequence 173, Application US/10176921			
Publication No. US2003002726A1			
GENERAL INFORMATION:			
APPLICANT: Baker, Kevin P.			
APPLICANT: Beresini, Maureen			
APPLICANT: Deforge, Laura			
APPLICANT: Desnoyers, Luc			
APPLICANT: Filvaroff, Ellen			
APPLICANT: Gao, Wei-Qiang			
APPLICANT: Gerritsen, Mary E.			
APPLICANT: Goddard, Audrey			
APPLICANT: Godowski, Paul J.			
APPLICANT: Gurney, Austin L.			
APPLICANT: Sherwood, Steven			
APPLICANT: Smith, Victoria			
APPLICANT: Stewart, Timothy A.			
APPLICANT: Tumas, Daniel			
APPLICANT: Watanabe, Colin K			
APPLICANT: Wood, William			
APPLICANT: Zhang, Zemin			
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
FILE REFERENCE: P330R1C288			
CURRENT APPLICATION NUMBER: US/10/176,921			
CURRENT FILING DATE: 2002-06-20			
Prior Application removed - See file Wrapper or Palm			
Number Of SEQ ID NOS: 550			



SEQ ID NO 173  
 LENGTH: 1210  
 TYPE: DNA  
 ORGANISM: Homo Sapien  
 US-10-176-921-173

## Alignment Scores:

Pred. No.:	4,2e-102	Length:	1210
Score:	848.00	Matches:	162
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-082-502-20 (1-162) x US-10-176-921-173 (1-1210)

```

QY 1 ArgArGserGlnAspLeuH1sCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
DB 603 CGGAGGAGCCAGGATCTCCACTGTGGAGCATGCGAGGCTGTGGTGGATGAACTAGAAATGG 662
QY 21 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40
DB 663 GAATATGCCAGGTGGACCCCAAGAAAGACCATTCAGATGGGATCTTCCGATCAATCCA 722
QY 41 AspGlySerGlnSerValGluValProTyraAlaArgSerGluAlaHisLeuThrGlu 60
DB 723 GATGGCAGCCAGTCAGTGGTGGAGGTGCTTATGCCCCGCTCAGAGGCCACCTCAGAGAG 782
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 80
DB 783 CTCTGTGAGGAGATATGTGACCGGATGAAGAGATATGGGAACAGATTCCTTCCACC 842
QY 81 HisArgLysAsnTyraValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 100
DB 843 CATCGCAAGAACTACGTAGCTGTAGTGGCCCGGAATGGAAATCCACTGAGACTGA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 903 CAAGGCATCCGAATCGACATGATATTACGGCACCCCTCAAGTTTCCGTGTAGACACTT 962
QY 121 ValGluGluTyrglyAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
DB 963 GTGGAGGAATACGAGGATGAACTCATTTCTTTCCCGAGAGGCTGACAAATGTTAAA 1022
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
DB 1023 GACAAACTTTCAGTAGCGAAGACAGATCTTTGTGACATGCCCTGCACATATGCATGAT 1082
QY 161 GluLeu 162
DB 1083 GAGCTA 1088

```

## RESULT 9

US-10-137-865-173  
 Sequence 173, Application US/10137865  
 Publication No. US20030032155A1

## GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
 APPLICANT: Beresini, Maureen  
 APPLICANT: Deforge, Laura  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Filvaroff, Ellen  
 APPLICANT: Gao, Wei-Qiang  
 APPLICANT: Gerritsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Goddard, Paul J.  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Sherwood, Steven  
 APPLICANT: Smith, Victoria  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K  
 APPLICANT: Wood, William

APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 FILE OF INVENTION: ACIDS ENCODING THE SAME  
 FILE REFERENCE: P33030R1C154  
 CURRENT APPLICATION NUMBER: US/10/137, 865  
 PRIOR APPLICATION REMOVED - See Palm or File Wrapper  
 NUMBER OF SEQ ID NOS: 550  
 SEQ ID NO 173

LENGTH: 1210  
 TYPE: DNA  
 ORGANISM: Homo Sapien  
 US-10-137-865-173

## Alignment Scores:

Pred. No.:	4,2e-102	Length:	1210
Score:	848.00	Matches:	162
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-082-502-20 (1-162) x US-10-137-865-173 (1-1210)

```

QY 1 ArgArGserGlnAspLeuH1sCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
DB 603 CGGAGGAGCCAGGATCTCCACTGTGGAGCATGCGAGGCTGTGGTGGATGAACTAGAAATGG 662
QY 21 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40
DB 663 GAATATGCCAGGTGGACCCCAAGAAAGACCATTCAGATGGGATCTTCCGATCAATCCA 722
QY 41 AspGlySerGlnSerValGluValProTyraAlaArgSerGluAlaHisLeuThrGlu 60
DB 723 GATGGCAGCCAGTCAGTGGTGGAGGTGCTTATGCCCCGCTCAGAGGCCACCTCAGAGAG 782
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 80
DB 783 CTCTGTGAGGAGATATGTGACCGGATGAAGAGATATGGGAACAGATTCCTTCCACC 842
QY 81 HisArgLysAsnTyraValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 100
DB 843 CATCGCAAGAACTACGTAGCTGTAGTGGCCCGGAATGGAAATCCACTGAGACTGA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 903 CAAGGCATCCGAATCGACATGATATTACGGCACCCCTCAAGTTTCCGTGTAGACACTT 962
QY 121 ValGluGluTyrglyAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
DB 963 GTGGAGGAATACGAGGATGAACTCATTTCTTTCCCGAGAGGCTGACAAATGTTAAA 1022
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
DB 1023 GACAAACTTTCAGTAGCGAAGACAGATCTTTGTGACATGCCCTGCACATATGCATGAT 1082
QY 161 GluLeu 162
DB 1083 GAGCTA 1088

```

## RESULT 10

US-10-140-474-173  
 Sequence 173, Application US/10140474  
 Publication No. US20030032156A1

## GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
 APPLICANT: Beresini, Maureen  
 APPLICANT: Deforge, Laura  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Filvaroff, Ellen  
 APPLICANT: Gao, Wei-Qiang  
 APPLICANT: Gerritsen, Mary E.  
 APPLICANT: Goddard, Audrey

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Saplen
US-10-140-474-173

Alignment Scores:
Pred. No.: 4.2e-102      Length: 1210
Score: 848.00           Matches: 162
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 9                      Gaps: 0

US-10-082-502-20 (1-162) x US-10-140-474-173 (1-1210)

QY 1 ArgArgSerGlnAspLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuTrp 20
DB 603 CGAGAGACCCAGATCTCCACTGTGAGCATGCGAGGCTGTGGTGAATAGAAATGG 662
QY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetLysSerPheArgIleAsnPro 40
DB 663 GAATATGCCAGGTGGAGCCCAAGAACCATTCAGATGGATCTTCCGATCAATCCA 722
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
DB 723 GATGGACGCCAGTCACTGCTGAGAGTGCCCTTATGCCCCGCTCAAGGCCCTCAGAGAG 782
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
DB 783 CTCTGAGAGAGATATGTGACCGGATGAAGAGATGAGGAGAACAGATTCCTCCACC 842
QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerGluLeuAspLeu 100
DB 843 CATCGCAAGAACTACGACTGTAGTGGGCGGAAATGGAATCCAGTGAATCGAACCTA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 903 CAAGGCATCCGATTCAGCTAGATATTAGGGCACCCTCAAGTTTGCTGTAGAGCATTT 962
QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
DB 963 GTGAGAGAAATACAGATGAACTGAAATCTTTTCCGAGAGGCTGACAAATGTTAAA 1022
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisasp 160
DB 1023 GACAAACTTTGACGATGAAGCAACAGATCTTTGTGACCATGCCCTGCACATATGCATGAT 1082
QY 161 GluLeu 162
DB 1083 GACCTA 1088

RESULT 11
US-10-142-431-173
; Sequence 173, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Saplen
US-10-142-431-173

Alignment Scores:
Pred. No.: 4.2e-102      Length: 1210
Score: 848.00           Matches: 162
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 9                      Gaps: 0

US-10-082-502-20 (1-162) x US-10-142-431-173 (1-1210)

QY 1 ArgArgSerGlnAspLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuTrp 20
DB 603 CGAGAGACCCAGATCTCCACTGTGAGCATGCGAGGCTGTGGTGAATAGAAATGG 662
QY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetLysSerPheArgIleAsnPro 40
DB 663 GAATATGCCAGGTGGAGCCCAAGAACCATTCAGATGGATCTTCCGATCAATCCA 722
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
DB 723 GATGGACGCCAGTCACTGCTGAGAGTGCCCTTATGCCCCGCTCAAGGCCCTCAGAGAG 782
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
DB 783 CTCTGAGAGAGATATGTGACCGGATGAAGAGATGAGGAGAACAGATTCCTCCACC 842
QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerGluLeuAspLeu 100
DB 843 CATCGCAAGAACTACGACTGTAGTGGGCGGAAATGGAATCCAGTGAATCGAACCTA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 903 CAAGGCATCCGATTCAGCTAGATATTAGGGCACCCTCAAGTTTGCTGTAGAGCATTT 962
QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
DB 963 GTGAGAGAAATACAGATGAACTGAAATCTTTTCCGAGAGGCTGACAAATGTTAAA 1022
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisasp 160
DB 1023 GACAAACTTTGACGATGAAGCAACAGATCTTTGTGACCATGCCCTGCACATATGCATGAT 1082
QY 161 GluLeu 162
```

Db 1083 GAGCTA 1088  
|||||  
RESULT 12  
US-10-143-114-173  
; Sequence 173, Application US/10143114  
; Publication No. US20030036180A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Collin K  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C211  
; CURRENT APPLICATION NUMBER: US/10/143,114  
; PRIOR FILING DATE: 2002-05-09  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-143-114-173  
  
Alignment Scores:  
Pred. No.: 4.2e-102 Length: 1210  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 9  
US-10-082-502-20 (1-162) x US-10-143-114-173 (1-1210)  
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
Db 603 CGAGAGACCGAGGATCTCCAGTGTGAGCATGCAAGGCTCTGTGATGAATCAATG 662  
QY 21 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
Db 663 GAAATTGCCCGAGTGGACCCCAAGAGACCATTCAGATGGATCTTCCGATCAATCA 722  
QY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
Db 723 GATGGACGACGATGAGTGTGAGTGGCTTATGCCCTCAGAGGCCACCTCAGAG 782  
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
Db 783 CTGCTGGAGAGATATGTGACCGGATGAAAGAGTATGGGGAACAGATTGATCTCCACC 842  
QY 81 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyGlnSerSerGluLeuAspLeu 100  
Db 843 CATCCCAAGAACTACTAGTACGTACTATGGCCCGAATGGAATCCAGTGAACCTGACCTA 902  
QY 101 GlnGlyIleArgGlyLeuAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
Db 903 CAAGGCATCCGCAATCGATCAATATATAGCGCACCTCAAGTTTGCTGTGAGAGCAT 962  
QY 121 ValGluGluTyrGluAspGluLeuLeuIleGluPheSerArgGluAlaAspAsnValLys 140

Db 963 GTGAGGAATACGACGATGACATTCATTCTTTCCCGAGAGCTGACAACTGTAA 1022  
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
Db 1023 GACAACCTTTGCATGAAGCAACAGACATCTTGTGACCAATGCCCTGCACATTCGATGAT 1082  
QY 161 GluLeu 162  
Db 1083 GAGCTA 1088  
|||||  
RESULT 13  
US-10-140-002-173  
; Sequence 173, Application US/10140002  
; Publication No. US20030037623A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Collin K  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C39  
; CURRENT APPLICATION NUMBER: US/10/140,002  
; PRIOR FILING DATE: 2002-05-06  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-140-002-173  
  
Alignment Scores:  
Pred. No.: 4.2e-102 Length: 1210  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
US-10-082-502-20 (1-162) x US-10-140-002-173 (1-1210)  
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
Db 603 CGAGAGACCGAGATCTCCAGTGTGAGCATGCAAGGCTCTGTGATGAATCAATG 662  
QY 21 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
Db 663 GAAATTGCCCGAGTGGACCCCAAGAGACCATTCAGATGGATCTTCCGATCAATCA 722  
QY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
Db 723 GATGGACGACGATGAGTGTGAGTGGCTTATGCCCTCAGAGGCCACCTCAGAG 782  
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
Db 783 CTGCTGGAGAGATATGTGACCGGATGAAAGAGTATGGGGAACAGATTGATCTCCACC 842  
QY 81 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyGlnSerSerGluLeuAspLeu 100

```
|||||
Db 843 CATGCGAAGACTACGTACGTGTAGTGGCCGGAATGGAATCCATGCACTGGACCTA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuIysPheAlaCysGluSerIle 120
Db 903 CAGGCACTTCGAGATTCGATCAGATATTAGCGGACCCCTCAAGTTGGCTGTGAGAGCAT 962
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
Db 963 GTGGAGGAATACGAGATGCAATCATTCATTCCTTTCCGAGAGCGCTGACAAATGTTAAA 1022
QY 141 AspIysLeuCysSerIysAspGlyThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
Db 1023 GACAACTTTGCAAGTAAAGCAACAGATCTTTGTGACCATCCCTGCACATATGCGATGAT 1082
QY 161 GluLeu 162
Db 1083 GAGCTA 1088

RESULT 14
US-10-142-419-173
; Sequence 173, Application US/10142419
; Publication No. US20030044945A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Zhang, Zemin
; APPLICANT: Wood, William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C244
; CURRENT APPLICATION NUMBER: US/10/142,419
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Saplen
US-10-142-419-173

Alignment Scores:
Pred. No.: 4,2e-102 Length: 1210
Score: 848.00 Matches: 162
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-082-502-20 (1-162) x US-10-142-419-173 (1-1210)
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyr 20
Db 603 CGAGAGAGCCAGGATCTCCCTGTGAGCATGCAAGGCTCTGTGTGATGAACTAGAAATGG 662
QY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
Db 663 GAAATTCGCCAGGTGAGCCCAAGAGACCAATTCAGATGGAGATCTTCCGAGCAATCA 722
QY 41 AspgIysSerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
```

```
|||||
Db 723 GATGCGAGCCAGTCAGTGAAGTGGAGGTCCCTTATGCCCGCTCAGAGGCCCACTCACAGAG 782
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
Db 783 CTGTGGAGAGGATATATGTGACCGGATAGAGAGATATGGGAAACAGATTGCTCTCCACC 842
QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerGluLeuAspLeu 100
Db 843 CATGCGAAGACTACGTACGTGTAGTGGCCGGAATGGAATCCATGCACTGGACCTA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuIysPheAlaCysGluSerIle 120
Db 903 CAGGCACTTCGAGATTCGATCAGATATTAGCGGACCCCTCAAGTTGGCTGTGAGAGCAT 962
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
Db 963 GTGGAGGAATACGAGATGCAATCATTCATTCCTTTCCGAGAGCGCTGACAAATGTTAAA 1022
QY 141 AspIysLeuCysSerIysAspGlyThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
Db 1023 GACAACTTTGCAAGTAAAGCAACAGATCTTTGTGACCATCCCTGCACATATGCGATGAT 1082
QY 161 GluLeu 162
Db 1083 GAGCTA 1088

RESULT 15
US-10-123-262-173
; Sequence 173, Application US/10123262
; Publication No. US20030049816A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C38
; CURRENT APPLICATION NUMBER: US/10/123,262
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Saplen
US-10-123-262-173

Alignment Scores:
Pred. No.: 4,2e-102 Length: 1210
Score: 848.00 Matches: 162
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-082-502-20 (1-162) x US-10-123-262-173 (1-1210)
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyr 20
```

```
Db 603 CGGAGGACCCAGGATCTCCACTGTGGACCATCGAGGCTCTGGTGATGAACTAGAAATGG 662
QY 21 GLuIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
Db 663 GAAATGGCCAGGTGGACCCCAAGAGACCAATTCAGATGGGATCTTCCGGATCAATCCA 722
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
Db 723 GATGGCAGCCAGTCAGTGTGGAGGTGCTTATGCCCTCAGAGGCCACCTCACAGAG 782
QY 61 LeuGluGluValIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
Db 783 CTGCTGGAGGAGATATGTGACCGGATGAAGGATGGGGAACAGATTGATCTTCCACC 842
QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerSerGluLeuAspLeu 100
Db 843 CATCCGCAAGAACTACGTACGTGTACTGGCCCGGAATGGAGAAATCCAGTGAATCGACCTA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
Db 903 CAAGGCATCCGAAATCGACTCAGATATATAGCGGCACCCCTCAAGTTGCGTGTGAGACATT 962
QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
Db 963 GTGGAGGAATACGAGGATGAACCTCAATTGATCTTTCCCGAGAGGCTGACAAATGTAAA 1022
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
Db 1023 GACAACCTTTGCAGTAACGAAACAGATCTTGTGTGACCATGCCCTGCACATATCGCATGAT 1082
QY 161 GlnLeu 162
Db 1083 GAGCTA 1088
```

Search completed: March 31, 2003, 11:59:11  
Job time : 51.9767 secs

10

11

12

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:33:58 ; Search time 35.3198 Seconds  
(without alignments)  
1406.624 Million cell updates/sec

Title: US-10-082-502-20  
Sequence: 1 RRSODLHCAGCALVDELEW.....LCSTRDLCDAHLSHDEL 162

Scoring table: BLOSUM62  
Xgapop 10.0, Ygapext 0.5  
Xgapop 6.0, Ygapext 7.0  
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Command line parameters:  
-MODEL=frame\_plus\_p2n.model -DEV=xip  
-Q/cgpn2.1/USPTO.spool/US10082502/rnuna.24032003.135100.6616/app\_query.fasta.1.1308  
-DB=Issued\_Patents\_NA -ORF=fastab -SUFFIX=rm1 -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdl  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pcpt -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRR=US10082502.ecgn.1.131.0runat.24032003.135100.6616 -NCPU=6 -ICPU=3  
-MARN.TIMEOUT=30 -THREADS=1 -GAPOP=10 -XGAPEXT=0.5 -WAIT -LONGLOG -DRV.TIMEOUT=120  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA:  
1: /cgpn2.6/prodata/1/lna/5A.COMB.seq:\*  
2: /cgpn2.6/prodata/1/lna/5B.COMB.seq:\*  
3: /cgpn2.6/prodata/1/lna/6A.COMB.seq:\*  
4: /cgpn2.6/prodata/1/lna/6B.COMB.seq:\*  
5: /cgpn2.6/prodata/1/lna/6C.COMB.seq:\*  
6: /cgpn2.6/prodata/1/lna/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	15.2	744	4	US-09-163-285-3
2	129	15.2	1512	4	US-09-163-285-1
3	91.5	10.8	915	4	US-09-724-864-14
4	85	10.0	5181	1	US-08-257-073-10
5	81.5	9.6	1924	3	US-08-961-083-159
6	81.5	9.6	4394	2	US-08-750-152A-1
7	80.5	9.5	1216	4	US-09-071-035-3
8	80.5	9.5	1347	4	US-09-071-035-1
9	80.5	9.5	3752	4	US-08-961-527-208
10	77.5	9.1	2508	2	US-08-850-993-1
11	76.5	9.0	4865	3	US-08-894-017-24
12	75	8.8	5558	4	US-08-961-527-103

13	74.5	8	833	2	US-08-837-029-1	Sequence 1, Appl1
14	74.5	8	3372	1	US-07-906-349A-1	Sequence 1, Appl1
15	74.5	8	3372	1	US-08-167-035-1	Sequence 1, Appl1
16	74.5	8	3372	1	US-08-167-035-1	Sequence 48, Appl1
17	74.5	8	3372	1	US-08-208-887A-1	Sequence 1, Appl1
18	74.5	8	3372	2	US-08-539-005-1	Sequence 1, Appl1
19	74.5	8	3372	2	US-08-539-005-1	Sequence 48, Appl1
20	74.5	8	3372	3	US-09-344-521-1	Sequence 1, Appl1
21	74.5	8	3372	4	US-09-280-598-1	Sequence 1, Appl1
22	74.5	8	5177	6	5352450-1	Patent No. 5352450
23	74	8.7	5661	4	US-08-938-105-2	Sequence 2, Appl1
24	73.5	8.7	1776	1	US-08-522-229B-1	Sequence 1, Appl1
25	73.5	8.7	1776	2	US-09-102-433-1	Sequence 1, Appl1
26	73	8.6	2012	2	US-08-484-200-3	Sequence 3, Appl1
27	73	8.6	4131	1	US-07-872-644-38	Sequence 38, Appl1
28	73	8.6	4131	1	US-08-297-494-38	Sequence 38, Appl1
29	73	8.6	4131	1	US-08-297-510-38	Sequence 38, Appl1
30	73	8.6	4131	1	US-08-479-532-38	Sequence 38, Appl1
31	73	8.6	4131	1	US-08-455-526-38	Sequence 38, Appl1
32	73	8.6	4131	1	US-08-455-526-38	Sequence 38, Appl1
33	73	8.6	4131	3	US-09-139-491-38	Sequence 38, Appl1
34	73	8.6	4131	5	PCT-US92-03222-38	Sequence 38, Appl1
35	72.5	8.5	4371	1	US-08-803-973-1	Sequence 1, Appl1
36	72.5	8.5	4371	1	US-08-803-973-1	Sequence 1, Appl1
37	72	8.5	1640	4	US-08-676-444-41	Sequence 41, Appl1
38	71.5	8.4	676	4	US-08-998-416-1130	Sequence 1130, Ap
39	71.5	8.4	724	4	US-08-998-416-928	Sequence 928, Ap
40	71.5	8.4	41708	4	US-09-470-512A-3	Sequence 3, Appl1
41	71	8.4	2073	4	US-09-134-001C-2101	Sequence 2101, Ap
42	71	8.4	2160	2	US-08-840-236-2	Sequence 2, Appl1
43	71	8.4	2160	2	US-08-840-236-5	Sequence 2, Appl1
44	71	8.4	2160	2	US-08-505-448A-2	Sequence 2, Appl1
45	71	8.4	2160	2	US-08-505-448A-5	Sequence 5, Appl1

## ALIGNMENTS

RESULT 1  
US-09-163-285-3  
Sequence 3, Application US/09163285  
Patent No. 6204013  
GENERAL INFORMATION:  
APPLICANT: Rhododoust, Mehran  
TITLE OF INVENTION: NOVEL MSP-5 PROTEIN AND NUCLEIC ACID MOLECULES  
NUMBER OF SEQUENCES: 4  
TITLE OF INVENTION: AND USES THEREOF  
ADDRESS: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/163,285  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/090,398  
FILING DATE: June 24, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandigouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MNI-049  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 3:





Alignment Scores:	
Pred. No.:	0.0112
Score:	91.50
Percent Similarity:	42.07%
Best Local Similarity:	24.39%
Query Match:	10.75%
DB:	4
	Indels:
	Gaps:
	8
	Length:
	915
	Matches:
	40
	Conservative:
	29
	Mismatches:
	66
	Indels:
	29
	Gaps:
	8

OY		b	LeuHisCysGluValAspArgLalaValAlaPheLeuGluIutPrpGluIleAlaGlnVal	25
Db		158	CTGGCGTCGCATGCCCTCCGGCGCTGTGGCCCTTCCAGATGGGGCAACCTCTGGCCAAAAGA	217
OY		26	AspProLysIsthrIleGlnMetGlySerPheArgIleAsnProAspGlySerGlnSer	45
Db		218	GAGCGTAATAATCTCAC-----ACTCCAGACGCCAGT---GGA	250
OY		46	ValValGlaValProTyrLalaIatSerGluAlaHisLeuThrGluLeuGluGluIle	65
Db		251	TTCGAGGAGCTG-----AGTGAAATCCACGTCACAGATGTCTCGAACCAACC	298
OY		66	CysAspArg---MetylsGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsn	84
Db		299	TGCCTTCAGACACTGGCCAGTCTATTGA-----GTTCAATGAAGTAGAAC	340

QY	85	TyValArgAlaValGlyArganglylSerSercIuEuaSpLeunginglyllearg	104
	:::	::	
Db	341	CAGATGAACGGCTCAACGGGCCAAGACTTACCAAGGGCCAAAG-----CCAAGA	391
	:::	::	
QY	105	IleaspSerapIlleserlg-----ThreunysPhealacysglusler	119
		:::	

[illegible]

```

QY      159 H1SAspGIUeu 162
          :::|||||
Db      572 AGAGAGAGAGCTT 583

```

RESULT 4  
US-08-257-073-10  
; Sequence 10, Application US/08257073  
Dataset Name: E7C6107

```

GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: de Talsne, Charles
APPLICANT: Tine, John A.
TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE

```

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,783  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/852,305  
FILING DATE: 18-MAR-1992  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER: US 07/672,183  
FILING DATE: 20-MAR-1991  
ATTORNEY/AGENT INFORMATION:

```

: REGISTRATION NUMBER: 25,506
: REFERENCE/DOCKET NUMBER: 454310-2570
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 840-9333
: TELEFAX: (212) 840-0712
: TELEX: 425066 CURTMS
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5181 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
US-08-257-073-10

```

US-08-257-073-10

Alignment Scores:	
Pred. No.:	0.956
Score:	85.00
Percent Similarity:	45.31%
Best Local Similarity:	22.93%
Query Match:	10.02%
DB:	1
	7

US-10-082-502-20 (1-162) x US-08-257-073-10 (1-5181)

16 ASPGIUENGLIWPGLURIEALAGINVAASPProLysThrIleGlnmetGlySer 35  
 ||| ||||| ||| ::| ||| ::| ::|  
 Db 1192 GATCCACACTTGAATTTAGAAATTATTATTAAAGGAAAAAATAAAAAAGTTGATGTAAACACT 125

Db 1252 AATCACAAGATCTACGAATCTGTCAATATCCAAAGTTCTTATCCAAATGTAAT 1311

Db 1312 G T A T A T C C T T T A C C A C T G A T T C A T T C A T T A G C T G C A G A T A T G A T A A A A T 1371

Db 1372 TCATATGCTGATTTAATGATCCGTGATACTAAGAATAAATTAATGAAATAATATTATACCA 1431

Db 1432 GATAATAGGAGAAATAATATTCATTATATACATTAAAAACMAATTGATTTAGAGAA 1491

QY 103 ILIARGILEASPSEASPILIESEGLYTHRLEULYSPHEALCYSGIUSERILEVALGLU 122

DB 1492 AAAAAATTAATAC-----ACAAAAGACAAATAAAAAATTAATTGAA 1536  
QY 123 GATGTCGlu-----AspGluLeuIleGlu---PhePheSerArgGlu 135  
DB 1537 GATTATGAAAGCAAAAAAGCATTTATGAGATTAAGTCAAAAAATTTTGAATGAAA 1596  
QY 136 AlaAspAsn-----ValLysAspLysLeuGlySerLysArg 147  
DB 1597 TTTAATTAATTAATTTTGCACAAAGATGCTGATTAATAATATTTCAGTCACAA 1647  
RESULT 5  
US-08-961-083-159  
Sequence 159, Application US/08961083  
Patent No. 6159469  
GENERAL INFORMATION:  
APPLICANT: Choi et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,083  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 159:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1924 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-961-083-159  
Alignment Scores:  
Pred. No.: 0.655 Length: 1924  
Score: 81.50 Matches: 32  
Percent Similarity: 39.738 Conservativeness: 26  
Best Local Similarity: 21.928 Mismatches: 65  
Query Match: 9.61% Indels: 23  
DB: 3 Gaps: 4  
US-10-082-502-20 (1-162) x US-08-961-083-159 (1-1924)  
QY 15 ValAspGluLeuGluTyrGluIleAla-----GlnValAspProLysLysThr 30  
DB 575 ATTCTACTTGGAGACAAAGTTGCTACTGCTCAACATCAAGTAGATTATTTGAAAAA 634  
QY 31 IleGlnMetGlySerPheArgIleAsnProAspGlySerGlnSerValValGluValPro 50  
DB 635 CTCTCTTGCTGGTGGC-----GATCCTGATGATGGACAGAGATTATAGAAAGCTAAA 685  
QY 51 TTTAAlaArgSerGluAlaHisLeuThrGluLeuLeuGluGluIleGlyAspArgMetLys 70

DB 686 TTTAAAAAGAGAGAGCTGAGCTTAACGCTTAACAGCTGATAGCAAAAAACAAACA 745  
QY 71 GATGTCGluGluGlnIleAspProSerThrHisArgLysAsnTyrValArgValIleGly 90  
DB 746 GAACCTTAAGAAACTCTTGACAGCGCTGATCCGTAAGCTCAAGCTCAGATGATGAT 805  
QY 91 ArgAsnGlyGlySerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAspIleSer 110  
DB 806 AAGAGACAGACAGAGAGCTGATGATTAAGCTGATGATGATGATGATGATGATGAT 865  
QY 111 G1ThrLeuLysPheAlaCysGluSerIleValGluGluTyrGluAspGluLeuIleGlu 130  
DB 866 GATTTA-----GAAAAAGAAATTAAGTAACTTGAACCTTGA 895  
QY 131 PhePheSerArgGluAlaAspAsn-----ValLysAspLysLeuGly 144  
DB 896 ATATTACTTGAGAGCGCTGATNCTGAAGATGATGATGCTGCTCTCAATAATTAATGCT 955  
QY 145 SerLysArgThrAspLeu 150  
DB 956 ACTAAAAAGCTGAATTG 973  
RESULT 6  
US-08-750-152A-1  
Sequence 1, Application US/08750152A  
Patent No. 5977331  
GENERAL INFORMATION:  
APPLICANT: ASAKURA, YOKO  
APPLICANT: KIMURA, ETICHIRO  
APPLICANT: ABE, CHIZU  
APPLICANT: KAWAHARA, YOSHIO  
APPLICANT: KAWAMATSU, TSUYOSHI  
TITLE OF INVENTION: ALPHA-KETOGLUTARATE DEHYDROGENASE GENE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT, P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22152  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/750,152A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: NORMAN F. OBLON  
REGISTRATION NUMBER: 24,618  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4394 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Brevibacterium lactofermentum  
STRAIN: ATCC13869  
FEATURE:  
NAME/KEY: CDS







QY 88 ValValGlyArgAnGlyGlySerSerGluLeuAspLeuGlnGlyIleArgIle----- 105  
Db 380 CTGATC--TTGATCAGACAGCAAAAGATGCTAGTCTGCT-GGTCGAATGTTGTCAA 435  
QY 106 AspSerAspIleSer---GlyThrLeuLysPheAlaCysGlySerIleValGluGlyTyr 124  
Db 436 GATGCCATTTATTAAGAACTGTTAAACAGCTCAAGACAGTC-----CAAAA 489  
QY 125 GluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLysAspLysLeuCys 144  
Db 490 GAAAGTGAATTAAGAAAGATACACAAACAGCTGAGATATTAAGAACAGACAGAT 549  
QY 145 SerLysArgThrAspLeuCysAspHisAlaLeuHisIleSer 158  
Db 550 CAATATTAATCGATGTAGCTGCTCATGAGCAGAAAGTCTGCT 591  
RESULT 12  
US-08-961-527-103  
Sequence 103, Application US/08961527  
Patent No. 6420135  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 391  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,527  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5558 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-961-527-103  
Alignment Scores:  
Pred. No.: 21.2 Length: 5558  
Score: 75.00 Matches: 36  
Percent Similarity: 37.85% Conservative: 31  
Best Local Similarity: 20.34% Mismatches: 62  
Query Match: 8.84% Indels: 48  
Gaps: 8  
US-10-082-502-20 (1-162) x US-08-961-527-103 (1-5558)  
QY 14 LeuValAspGluLeu-----GluTrpGluIleAlaGlnValAspProLys 28  
Db 2094 ATGATAGACAAATATTAATACGTCATTTGAGGAATGCAATATCTGCGAAATGAGAACAA 2153  
QY 29 LysThrIleGlnMetGlySerPheArgIleAsnProAspLysSerGlnSerValValGlu 48

Db 2154 GGCATTCTCATAGTTTGGAGATAGCCGTTGAGCAATCATATATGCTACATCCAAAGTA 2213  
QY 49 ValProTyraIlaArgSerGluAlaHisLeuThrGluLeuLeuGlnGluIleCysAspArg 68  
Db 2214 GTTGGTGGGCTCGCCCAACAGCAAGATATACAGCTTTTCAAGAGAAAGTTTGC----- 2267  
QY 69 MetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyraValArgVal 88  
Db 2268 CAAGAAGCACCTTGCTGTTGTTG-----CCGACCTGACAAAGAAAGTTATCATCCGATATC 2318  
QY 89 ValGlyArg-----AsnGlyGlySerSer-----GluLeuAspLeuGlnGly 102  
Db 2319 CARGCCAAAGTCCATTAACACCAAGACGCGCTACTTCTGTCATATCTTACATCCAAAGGT 2378  
QY 103 IleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlySerIleValGlu 122  
Db 2379 ATAGAAATC---ACGACTCTCCAGACGCTTGAAGTTCTCC----- 2417  
QY 123 GluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys----- 140  
Db 2418 -----TACTTCTTCATGAGACATCAACATCCGGCGCTC 2453  
QY 141 -----AspLysLeuCysSer 145  
Db 2454 TTCAGCCATTTTCATATAGAGGTCATTGATACGAGCTTCAGCTTGTGAAGATCATCAAA 2513  
QY 146 LysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAspGluLeu 162  
Db 2514 AGCCGTACGAGAAACATACACGACCGACGCTGCTTTCAGCAAGACAGAGTG 2564  
RESULT 13  
US-08-837-029-1  
Sequence 1, Application US/08837029  
Patent No. 5945303  
GENERAL INFORMATION:  
APPLICANT: Wei et al.  
TITLE OF INVENTION: Human Hematopoietic - Specific Protein  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: MD  
COUNTRY: US  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/837,029  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US96/04930  
FILING DATE: 11-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PF268  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 833 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:

NAME/KEY: CDS  
 LOCATION: 42..608  
 FEATURE: sig\_peptide  
 LOCATION: 42..107  
 NAME/KEY: mat\_peptide  
 LOCATION: 108..608  
 US-08-837-029-1

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	1.6	833	34	29	75	23
Percent Similarity:	74.50					
Best Local Similarity:	39.13%					
Query Match:	21.12%					
	8.79%					

US-10-082-502-20 (1-162) x US-08-837-029-1 (1-833)

QY 6 LeuHlaCysGlyAlaCysArgAlaLeuValAspLeuLeuGluTrpGluIleAlaGluVal 25  
 DB 183 CTGGCGCTGTATGCTCTGACAGAGCTGTGCTTACCAGATGGCAAAATCTGGCAAGGCA 242  
 QY 26 AspProLysLysThrIleGlnMetGlySerPheArgIleAsnProAspGlySerGlnSer 45  
 DB 243 GAGACCAAACTGATACCTCA-----AACCTGGGGGGGGCGGAGAA 284  
 QY 46 ValValGluValProtyrAlaArgSerGluAlaHisLeuThrGluLeuGluGluIle 65  
 DB 285 CTGACGCAAGTTGGTCTAC-----ACGGATGCTCTGGACCGGCAAC 323  
 QY 66 CysAspArg---MetLysGluTrgIly---GluGlnIleAspProSerThrHisArgLys 83  
 DB 324 TGCTCCCGGAACTGGGAGAGCTACGAGATTCGAAAGTGAGAC----- 365  
 QY 84 AsnTyValArgValAlaGlyArgAsn---GlyGlnSerSerGluLeuAspLeuGlnGly 102  
 DB 366 CAAGTAAACGCTCTCAGAGGCCGAGACTAGGAGGGCCAGAGCCAGCATCAGCGTG 425  
 QY 103 IleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGlu 122  
 DB 426 ATGGTACACAGGGGGCCCTGGCTTACAGAGCTCTCCAGGACATTTGGACTTGGGG 485  
 QY 123 GluTrg---GluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLysAsp 141  
 DB 486 GAGTTGGAGAGACCAAGATGTATGAAAGCCCAACAAAGGCCGAGGGCTGTGGAGCA 545  
 QY 142 LysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAspGlu 161  
 DB 546 TTGCTATGTGGGGGAGCCCAAGGGGGCTGCTCAGAGAAAGTGTCAAGACAGAGAG 605  
 QY 162 Leu 162  
 DB 606 CTC 608  
 RESULT 14  
 US-07-906-349A-1  
 ; Sequence 1, Application US/07906349A  
 ; Patent No. 5434064  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Schlessinger, Joseph  
 ; APPLICANT: Skolnik, Edward Y.  
 ; APPLICANT: Margolis, Benjamin L.  
 ; TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR  
 ; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR ENDOCYTIC TYROSINE KINASES AND  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Broadway and Nelmark  
 ; STREET: 419 Seventh Street, N.W.  
 ; CITY: Washington  
 ; STATE: D.C.

COUNTRY: USA  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/906,349A  
 FILING DATE: 30-JUN-1992  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/643,237  
 FILING DATE: 18-JAN-1991  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-628-5197  
 TELEFAX: 202-737-3528  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3372 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; US-07-906-349A-1  
 Alignment Scores:  

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	12	3372	32	57	35	8
Percent Similarity:	74.50					
Best Local Similarity:	40.65%					
Query Match:	20.65%					
	8.79%					

US-10-082-502-20 (1-162) x US-07-906-349A-1 (1-3372)  
 QY 22 IleAlaGluValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnProAsp 41  
 DB 1279 CTAGCTCAGTATTAATCCCAAAATGGATGTGAATTA-----CTTATTCAGTA 1326  
 QY 42 GlySerGlnSerValValGluValProtyrAlaArgSerGluAlaHisLeuThrGluLeu 61  
 DB 1327 TCCAAA-----TACCAACAGATCAAGTTGTCAAAAGAAATAT 1365  
 QY 62 LeuGluGluIleCysAspArgMetLysGluTrgIlyGluGlnIleAspProSerThrHis 81  
 DB 1366 ATTGAAGCTGTAGGAAATAATTAATGATTAACACTGTTTCAAGAAAAAGTCA 1425  
 QY 82 Arg-----LysAsnTyValArgValAlaGlyArgAsnGlyGluSerSer 96  
 DB 1426 GAATATGATAGTATATGAAAGAAATATACCCGACCA-----TCCAG 1467  
 QY 97 GluLeuAspLeuGlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPhe--- 115  
 DB 1468 GAATCCAAATGAAGAGACAGCTATTCAGCA---TTTATGAAACCAATAAAATATTT 1524  
 QY 116 -----AlaCysGluSerIleValGluGluTrgIlyAspLeuLeuIleGluPheSer 133  
 DB 1525 GAGACACAGTCCCGAGCC---CAAGAGCGGTACAGCAAGATATCATAGAAAGTTAA 1581  
 QY 134 ArgGluAlaAspAsnValLys-----AspLysLeuCysSer 145  
 DB 1582 CGTAAAGCAATGAGAAAGAAATATCAAAAGATATATGATATATGATATGATGAGTCA 1641  
 QY 146 LysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 DB 1642 CGAATCAGTGAATTTATTTGACAGTAGAAGAAAGATTTGGAAGAAAGAC 1686  
 RESULT 15  
 US-08-167-035-1  
 ; Sequence 1, Application US/08167035  
 ; Patent No. 5618691  
 ; GENERAL INFORMATION:

APPLICANT: Schlessinger, Joseph  
APPLICANT: Skolnick, Edward Y.  
APPLICANT: Margolis, Benjamin L.  
TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR  
TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR ENKAROTIC TYROSINE  
TITLE OF INVENTION: KINASES AND NOVEL TARGET PROTEINS  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: 10036-2711  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/167,035  
FILING DATE: 16-DEC-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7683-062  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3372 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 43..2214  
US-08-167-035-1  
Alignment Scores:  
Pred. No.: 12 Length: 3372  
Score: 74.50 Matches: 32  
Percent Similarity: 40.65% Conservative: 31  
Best Local Similarity: 20.65% Mismatches: 57  
Query Match: 8.79% Indels: 35  
DB: 1 Gaps: 8  
US-10-082-502-20 (1-162) x US-08-167-035-1 (1-3372)  
QY 22 ILAALAGLVALASPPOLYSLVSTHILGLMETCLYSERPHARGILEASNPASP 41  
DB 1279 CTAGCTGAGTAAATCCCAATTGGATGGAATA-----CTTATCCAGTA 1326  
QY 42 GlycerInserValValGluValProTyrAlaArgSerGluAlaHisLeuThrGluLeu 61  
DB 1327 TCCAAA-----TACCACAGCATCAAGTGTCCAAAGAAAGATAT 1365  
QY 62 LeuGluGluIleCysAspArgMetLysGluTyrGluGluGluIleAspProSerThrHis 81  
DB 1366 ATTGAAGCTGTAGGAAAAAATTACATGATATACACTGCTTCAAGAAAAAGTCGA 1425  
QY 82 Arg-----LysAsnTyrValArgValValGluValArgAsnGlyGluSerSer 96  
DB 1426 GAATATGATATGATATATGTAAGATATATACCCGACA-----TCCGAG 1467  
QY 97 GluLeuAspLeuGlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPhe--- 115  
DB 1468 GAATATCAATGAAAGAGACGATATGAGCA---TTTAAATGAACCATTAATAATT 1524

QY 116 -----AlaCysGluSerIleValGluGluTyrGluAspGluLeuIleGluPhePheSer 133  
DB 1525 GAAGAACAGTGCACAGCC-----CAAGACGGTACAGCAAAAGATATGAAAGTTTAAA 1581  
QY 134 ArgGluAlaAspAsnValLys-----AspLysLeuCysSer 145  
DB 1582 CGTGAAGCAATGAGAAAGAAATACAAAGATTATGCAATATATGATTAAGTTGAAGCT 1641  
QY 146 LysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
DB 1642 GAATATCAATGAAATATGACAGTGAAGAAAGATTTGGAAGAGAC 1686

Search completed: March 31, 2003, 04:26:11  
Job time : 42.3198 secs





```

FT      /*tag- a
FT      /product- "Human transmembrane protein HP10390"
FT      /note- "No stop codon given in the specification"
XX
XX      MO9955862-A2.
XX
XX      04-NOV-1999.
XX
XX      27-APR-1999; 99WO-JP02226.
XX
XX      28-APR-1998; 98JP-0119395.
XX
XX      (SAGA ) SAGAMI CHEM RES CENT.
XX      (PROT-) PROTEGENE INC.
XX
XX      Kato S, Kimura T;
XX
XX      WPI: 2000-023358/02.
XX      P-PSDB; AAY52391.
XX
XX      Human proteins with transmembrane domains, involved in control of cell
XX      proliferation and differentiation, useful for treating e.g. cancer or
XX      inflammation
XX
XX      Claim 3; Page 88; 114pp; English.
XX
XX      This sequence represents the coding sequence of human CDNA clone
XX      HP10390 which encodes a 20 kD protein with one putative transmembrane
XX      domain in the N-terminus. The CDNA was isolated from a human stomach
XX      cancer cell line CDNA library. The protein has no homology with any
XX      known protein. The protein may be used to raise specific antibodies, as
XX      assay reagents, as diagnostic tissue markers, for the isolation of
XX      cognate receptors, ligands and binding proteins, and as biologically
XX      active agents. Nucleotides encoding the protein may be used as primers
XX      and probes or antisense molecules, and in gene therapy. Cells transformed
XX      with these nucleotides may be used to screen for agonists and antagonists
XX      which are potentially useful therapeutically.
XX
XX      Sequence 546 BP; 139 A; 126 C; 165 G; 116 T; 0 other;
XX
XX
XX      Alignment Scores:
XX      Pred. No.: 2,2e-94 Length: 546
XX      Score: 848.00 Matches: 162
XX      Percent Similarity: 100.00% Conservative: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      DB: 21 Gaps: 0
XX
XX      US-10-082-502-20 (1-162) x AAZ38326 (1-546)
XX
XX      1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyr 20
XX      |||||||
XX      61 CGGAGGAGCGAGATCTCCACTGTGGAGCATGCGGCTGTGGTGAATGGAATGG 120
XX      |||||||
XX      21 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheAsnPro 40
XX      |||||||
XX      121 GAAATTCGCCAGGGGAGCCCAAGAACACATTCAGATGGAGATCTTCCGATCAATCCA 180
XX      |||||||
XX      41 AspGlySerGlnSerValIleGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
XX      |||||||
XX      181 GATGGGACGACAGTCAAGTGGGAGGTGCTTATGCCCGCTCAAGAGCCCACTCAGAGAG 240
XX      |||||||
XX      61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
XX      |||||||
XX      241 CTCCTGGAGAGATATGTGACCGGATGAAGAGATGGGAAACAGATTGCTTCCACC 300
XX      |||||||
XX      81 HisArgLysAsnTyrValaLysValaGlyArgAsnGlyLysSerGluLeuAspLeu 100
XX      |||||||
XX      301 CATCGGAAGAACTACACTGTGAGTGGCCCGGAATGGAAATCCACTGACACTGACCTA 360
XX      |||||||
XX      101 GlnGlyIleArgLysSerSerAspIleSerGlyThrLeuLysPheAlaCysGluSerTle 120
XX      |||||||
XX      361 CAAGGCAATCGAATTCAGTCAATTAATTAAGCGGACCTCAAGTTTGCTGTGAGAGCATTT 420

```

```

XX      121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgIleAlaAspAsnValLys 140
XX      |||||||
XX      421 GTGAGAGAAATACGAGAGATGACTGATTAATCTTTCCGAGAGGCTGACAAATGTTAAA 480
XX      |||||||
XX      141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
XX      |||||||
XX      481 GACAAACTTTCAGTGAAGCGAAGATCTTTGTGACCATGCTGCACATATGCATGAT 540
XX      |||||||
XX      161 GluLeu 162
XX      |||||||
XX      541 GAGCTA 546
XX
XX      RESULT 2
XX      ID AAX06969 standard; cDNA; 806 BP.
XX      XX
XX      AC AAX06969;
XX      XX
XX      DT 10-MAY-1999 (first entry)
XX      XX
XX      DE Human secretory peptide-9 (Zs1g9) variant cDNA.
XX      XX
XX      KW Secretory peptide-9; Zs1g9; human; tumour marker; cancer; therapy;
XX      KW diagnosis; growth enhancer; variant; ss.
XX      XX
XX      OS Homo sapiens.
XX      XX
XX      FH Key Location/Qualifiers
XX      FT CDS 104..652
XX      FT sig_peptide /*tag- a
XX      FT 104..163
XX      FT mat_peptide /*tag- b
XX      FT 164..649
XX      FT /*tag- c
XX
XX      PN MO9901554-A1.
XX      XX
XX      PD 14-JAN-1999.
XX      XX
XX      PF 02-JUL-1998; 98WO-US13859.
XX      XX
XX      PR 17-JUN-1998; 98US-0099005.
XX      PR 03-JUL-1997; 97US-0051704.
XX      PR 03-JUL-1997; 97US-0888088.
XX      PR 19-MAY-1998; 98US-0081338.
XX      PR 19-MAY-1998; 98US-0085983.
XX      PR 17-JUN-1998; 98US-0089899.
XX      XX
XX      PA (ZYMO ) ZYMOGENETICS INC.
XX      XX
XX      PI Jaspers SR, Jellinek LJ, Shepard PO, Whitmore TE;
XX      XX
XX      DR WPI: 1999-106055/09.
XX      DR P-PSDB; AAM88474.
XX      XX
XX      PT New mammalian secretory peptide-9 (Zs1g9) - used as a growth
XX      PT enhancer for placenta, liver and heart, and as an indicator of
XX      PT cancer
XX      XX
XX      PS Claim 2; Page 73-74; 85pp; English.
XX      XX
XX      CC This CDNA clone encodes human secretory peptide-9, or Zs1g9,
XX      CC variant (see AAM88474). Zs1g9 (see also AAM88469) is overexpressed in
XX      CC human brain, liver, lung, oesophageal, stomach, colon, rectal,
XX      CC thyroid and lymphoma tumors. Thus, Zs1g can be used as an
XX      CC indicator for cancer. Zs1g9 cDNA was discovered in a placenta
XX      CC clone from a full-term pregnancy cDNA library which contained an
XX      CC expressed sequence tag (see AAX06971). The invention provides
XX      CC polynucleotides (see AAX06968-70) encoding Zs1g9 polypeptides (see
XX      CC AAM88469-77) including mature polypeptides, other processed forms,
XX      CC variants and mouse orthologues. The Zs1g9 gene, or probes derived
XX      CC from it, can be used to determine if Zs1g9 is present on chromosome

```

CC 12, and if a mutation has occurred. Antibodies raised against  
 CC Zs19g can be used as diagnostic agents to determine the presence of  
 CC Zs19g, and thus the presence of cancer. They can also be labelled  
 CC with radiolabels or fused with toxins and used to treat tumours  
 CC which overexpress Zs19g. Antisense nucleotides derived from Zs19g  
 CC cDNA can also be used to inhibit the growth of tumour cells. Zs19g  
 CC proteins can be used to enhance the growth or development of the  
 CC placenta, heart or liver.

XX  
 SQ Sequence 806 BP; 229 A; 183 C; 232 G; 160 T; 2 other;

#### Alignment Scores:

Pred. No.:	3.69e-94	Length:	806
Score:	848.00	Matches:	162
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	20	Gaps:	0

US-10-082-502-20 (1-162) x AA06969 (1-806)

```

OY 1  ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
    |||||
DB 164 CGAGAGAGCCAGAGATCTCCACTGTGAGCATGCAGGCGCTGTGTGATCAATGATG 223
OY 21  GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
    |||||
DB 224 GAATATGCCAGAGTGCAGCCCAAGAACATTCAGATGGATCTTCCGATCAATCCA 283
OY 41  AspGlySerGlnSerValValGluValProTyraIleArgSerGluAlaHisLeuThrGlu 60
    |||||
DB 284 GATGGCAGCCAGTCACTGTGTGAGGTGGCTTATGCCCTGACAGGCCACCTCACAGAG 343
OY 61  LeuLeuGluGluIleCysAspArgMetLysGluTyrglyGlnIleAspProSerThr 80
    |||||
DB 344 CTGCTGGAGAGATATGTGACCGCATGAAGATATGGGACATGATCTTCCTCCACC 403
OY 81  HisArgLysAsnTyraValArgValAlaGlyArgAsnGlyGluSerSerGluLeuAspLeu 100
    |||||
DB 404 CATCCCAAGACTACGTACGTACTGTGGCCCGAATGGAGATCCAGTCACTGACACTCA 463
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
    |||||
DB 464 CAGGCAATCCGATGACATGACATATATGCGGACCCCTCAATGTCGTGAGACAT 523
OY 121 ValGluGluTyrgluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
    |||||
DB 524 GTGAGAGATACGAGATGAATCACTGATGATCTTTTCCGAGAGGCTGACATGTAA 583
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
    |||||
DB 584 GACAAACTTTCAGTAAAGAGAACAGATCTTGTGACATGCCCTGCACATATGCATGAT 643
OY 161 GluLeu 162
    |||||
DB 644 GAGCTA 649
  
```

#### RESULT 3

AA08293  
 ID AA08293 standard; DNA; 806 BP.

XX AA08293;

XX 07-FEB-2000 (first entry)

DE Human Zs19g gene encoding secretory protein variant-4.

KW Secretory protein-9; Human Zs19g; chromosome 12q15 region; variant;  
 KW overexpression; antagonist; antibody; antisense nucleotide; tumour;  
 KW treatment; receptor; radio-label; fusion; polypeptide toxin; technique;  
 KW down-regulation; probe; diagnostic; therapeutic; cancer; brain; liver;  
 KW detection; stomach; lymphoma; alternative splicing; allelic variation;  
 KW silent mutation; ds.

```

XX XX Homo sapiens.
OS OS Synthetic.
XX XX Key
FH FH CDS
FT FT Location/Qualifiers
FT FT 104..652
FT FT /*tag= a
FT FT /product= "Zs19g secretory protein variant-4"
FT FT /note= "Overexpressed in tumours"
FT FT sig_peptide
FT FT 104..163
FT FT /*tag= b
FT FT 164..649
FT FT /*tag= c
FT FT /label= Mature_Zs19g_protein_variant-4

XX PF WO960405-A1.
XX PF 19-MAY-1999; 99WO-US11107.
XX PF 19-MAY-1999; 98US-0081183.
XX PR 19-MAY-1998; 98US-0081183.
XX PA (ZYMO ) ZYMOGENETICS INC.
XX PI Moore EE, Taft DW;
XX PI WPI: 2000-039447/03.
XX DR P-PSDB; AA15135.
XX XX
XX PT Detecting tumors using antibodies, antagonists and antisense
XX PT nucleotides to secretory protein-9 (Zs19g)
XX PS
XX PS Disclosure: Page 33-35; 45pp; English.

XX CC The present DNA sequence is a gene encoding the variant of the secretory
XX CC protein-9, Zs19g that arises due to alternative splicing, allelic
XX CC variation or silent mutations that result in amino acid changes. This
XX CC sequence is mapped to the human chromosome 12q15 region. It is
XX CC overexpressed in tumors. Antagonists, antibodies and antisense
XX CC nucleotides to Zs19g are useful for detecting and treating tumors. The
XX CC antagonist may be an antibody or receptor to Zs19g and it may be radio-
XX CC labeled or fused to a polypeptide toxin. It can be used for down
XX CC regulating the overexpression of Zs19g. The gene sequence can be used as
XX CC nucleic acid probes to detect RNA encoding Zs19g. The Zs19g sequence
XX CC facilitates improved diagnostic and therapeutic techniques for detecting
XX CC and treating cancers, especially of the brain, liver, stomach, lymphoma,
XX CC etc., at an early stage.

XX SQ Sequence 806 BP; 229 A; 183 C; 232 G; 160 T; 1 other;

XX Alignment Scores:
XX Pred. No.: 3.69e-94 Length: 806
XX Score: 848.00 Matches: 162
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 21 Gaps: 0

XX US-10-082-502-20 (1-162) x AA08293 (1-806)
OY 1  ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
    |||||
DB 164 CGAGAGAGCCAGAGATCTCCACTGTGAGCATGCAGGCGCTGTGTGATCAATGATG 223
OY 21  GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
    |||||
DB 224 GAATATGCCAGAGTGCAGCCCAAGAACATTCAGATGGATCTTCCGATCAATCCA 283
OY 41  AspGlySerGlnSerValValGluValProTyraIleArgSerGluAlaHisLeuThrGlu 60
    |||||
DB 284 GATGGCAGCCAGTCACTGTGTGAGGTGGCTTATGCCCTGACAGGCCACCTCACAGAG 343
  
```





CC purifying the antibody from the animal. The antibodies can be used  
CC as drugs, vaccines, diagnostic reagents and laboratory reagents. The  
CC present sequence encodes a polypeptide, designated HPI0390, which was  
CC used in the course of the invention.

XX Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other:

## Alignment Scores:

Pred. No.:	3,74e-94	Length:	814
Score:	848.00	Matches:	162
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-10-082-502-20 (1-162) x ABL41995 (1-814)

```
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTyr 20
   |||
DB 205 CGGAGGAGCCGAGATCTCCACTGTGGAGCATGCGGCTCTGTGGATGAACCTGGAATGG 264

QY 21 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40
   |||
DB 265 GAATTTGCCAGGTGGACCCCAAGAAACCATTCAGATGGATCTTCCGGATCAATCCA 324

QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
   |||
DB 325 GATGGCAGCCAGTCAGTGGTGGAGAGTCCCTTATGCCGCTCAGAGCCCACTCAGAGAG 384

QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGlnIleAspProSerThr 80
   |||
DB 385 CTGCTGGAGGAGATATGTGACCGGATGAAGAGATATGGGAAACGATTGATCTTCCACC 444

QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGlnSerSerGluLeuAspLeu 100
   |||
DB 445 CATGGCAAGAACTACGATACGTGTAGTGGCCGGAATGAGATCCATGAACTGGACCTA 504

QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
   |||
DB 505 CAAAGCATCCGAAATCGACTCGATATTAACGCGCACCTCAAGTTGGCTGTGAGAGCAT 564

QY 121 ValGluGluTyrGlnAspGluLeuIleGluPheSerArgIleValAspAsnValLys 140
   |||
DB 565 GTGAGAGAAATACGAGGATGAACATCATGAACTTTTCCGAGAGGCTGACAAATGTTAAA 624

QY 141 AspLysLeuGlySerLysArgThrAspLeuGlyAspHisAlaLeuHisIleSerHisAsp 160
   |||
DB 625 GACAAACTTTCAGTAAAGCAAGATCTTGTGACCATGCTCGACATATCCATGAT 684

QY 161 GluLeu 162
   |||
DB 685 GAGCTA 690

RESULT 7
ABK09772 standard: cDNA; 814 BP.
XX ID ABK09772:
XX AC ABK09772:
XX 14-MAR-2002 (first entry)
XX DE Human ovarian tumour protein encoding cDNA #305.
XX KW Human; ovarian tumour protein; cancer; cytostatic; immunostimulant; ss;
XX gene therapy; CD4+ T cell; CD8+ T cell; PCR primer.
XX OS Homo sapiens.
XX PN WO200190154-A2.
XX PD 29-NOV-2001.
XX PF 23-MAY-2001; 2001WO-US16895.
```

XX 24-MAY-2000; 2000US-207107P.  
PR 13-JUN-2000; 2000US-211457P.  
PR 21-JUN-2000; 2000US-213673P.  
PR 03-AUG-2000; 2000US-223288P.  
PR 01-MAR-2001; 2001US-272790P.

XX (CORI-) CORIXA CORP.

XX Xu J, Mitcham JL, Harlocker SL, Dillon DC, Secrist H, Lodes MJ;  
PI Algate PA, Fling SP, Mannion J, Benson DR, Carter D;

DR WPI; 2002-097641/13.

PT New isolated polynucleotide encoding polypeptide comprising portion of  
PT ovarian tumour protein, useful for detection, diagnosis and therapy of  
PT human ovarian cancer

PS Claim 1; Page 253-254; 285pp; English.

CC The invention relates to an isolated polynucleotide encoding a  
CC polypeptide comprising a portion of an ovarian tumour protein. The  
CC sequences of the invention are useful for stimulating an immune response  
CC and for treating ovarian cancer in a patient. An antigen presenting cell  
CC that expresses the sequences is useful for treating ovarian cancer by  
CC incubating CD4+ and/or CD8+ T cells isolated from a patient. The T cells  
CC can then be proliferated and administered to the patient to inhibit the  
CC development of cancer. The DNA sequences are useful as probes or primers  
CC for nucleic acid hybridisation, to direct expression of a polypeptide in  
CC appropriate host cells, detecting the presence of a cancer in a patient  
CC involves obtaining a biological sample from the patient, contacting the  
CC biological sample with an agent that binds to the protein, detecting the  
CC amount of protein that binds to the agent, comparing the amount of  
CC protein to a predetermined cut-off value and determining the presence of  
CC cancer. Sequences ABK09464-ABK09802 represent PCR primers and cDNA  
CC molecules encoding ovarian tumour proteins of the invention.

XX Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other:

## Alignment Scores:

Pred. No.:	3,74e-94	Length:	814
Score:	848.00	Matches:	162
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-10-082-502-20 (1-162) x ABK09772 (1-814)

```
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTyr 20
   |||
DB 205 CGGAGGAGCCGAGATCTCCACTGTGGAGCATGCGGCTCTGTGGATGAACCTGGAATGG 264

QY 21 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40
   |||
DB 265 GAATTTGCCAGGTGGACCCCAAGAAACCATTCAGATGGATCTTCCGGATCAATCCA 324

QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
   |||
DB 325 GATGGCAGCCAGTCAGTGGTGGAGAGTCCCTTATGCCGCTCAGAGCCCACTCAGAGAG 384

QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGlnIleAspProSerThr 80
   |||
DB 385 CTGCTGGAGGAGATATGTGACCGGATGAAGAGATATGGGAAACGATTGATCTTCCACC 444

QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGlnSerSerGluLeuAspLeu 100
   |||
DB 445 CATGGCAAGAACTACGATACGTGTAGTGGCCGGAATGAGATCCATGAACTGGACCTA 504

QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
   |||
DB 505 CAAAGCATCCGAAATCGACTCGATATTAACGCGCACCTCAAGTTGGCTGTGAGAGCAT 564
```

OY 121 VALGLUGLUTYRGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValIys 140  
 DB 565 GTGGAGGAAATACGAGATCACTGTAATCTTTCCCGAGAGCTGCACATGTATAA 624  
 OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 DB 635 GACAAACTTTGGCAGTAAAGCAAGATCTTTGTGACACATGCCCTGCACATTCGACATGAT 684  
 OY 161 GluLeu 162  
 DB 685 GAGCTA 690  
 RESULT 8  
 AAX22112  
 ID AAX22112 standard; DNA; 1180 BP.  
 AC AAX22112;  
 XX  
 DT 18-MAY-1999 (first entry)  
 XX  
 DE Human secreted protein gene 2 clone H2MBB56.  
 XX  
 KW Human: secreted protein; gene therapy; protein therapy; tissue; cancer;  
 KW tumour; neurodegenerative disorder; leukaemia; autoimmune disease; AIDS;  
 KW developmental abnormality; foetal deficiency; Alzheimer's disease;  
 KW cognitive disorder; schizophrenia; immunological disorder; mood disorder;  
 KW immune deficiency disease; respiratory disorder; arthritis; skeletal;  
 KW haematopoietic disorder; neural; osteoporosis; metabolic disorders;  
 KW cardiovascular; endocrine; gastrointestinal; asthma; diagnosis; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PM MO9901020-A2.  
 XX  
 PD 14-JAN-1999.  
 XX  
 PF 30-JUN-1998; 98WO-US13608.  
 XX  
 PR 12-SEP-1997; 97US-0058663.  
 PR 01-JUL-1997; 97US-0051381.  
 PR 01-JUL-1997; 97US-0051480.  
 PR 12-SEP-1997; 97US-0056598.  
 XX  
 PA (HUMA-) HUMAN GENOME SCT INC.  
 XX  
 PI Carter KC, Endress GA, Feng P, Rosen CA, Ruben SM;  
 XX  
 DR WPI: 1999-105683/09.  
 DR P-PSDB: AAY01136, AAY01162, AAY01163.  
 XX  
 PT New isolated human genes and the secreted polypeptides they encode -  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, immune deficiency diseases or blood  
 PT disorders  
 XX  
 PS Claim 4: Page 117: 179pp; English.  
 XX  
 CC The invention relates to nucleic acid sequences (AAX22111 to AAX22134)  
 CC encoding human secreted proteins (AAY01135 to AAY01158). The secreted  
 CC protein gene sequences are deposited with the ATCC under deposit number  
 CC ATCC 209118. Host cells comprising recombinant vectors containing the  
 CC nucleic acid sequences are used for the recombinant production of the  
 CC secreted proteins. The polynucleotide and amino acid sequences are useful  
 CC for are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Pathological conditions can  
 CC be also diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new  
 CC polynucleotides. Specific uses are described for each of the  
 CC polynucleotides, based on which tissues they are most highly expressed  
 CC in, and include developing products for the diagnosis or treatment of  
 CC cancer, tumours, developmental abnormalities and foetal deficiencies,  
 CC autoimmune diseases, lymphomas, Alzheimer's and cognitive disorders,  
 CC schizophrenia, immunological disorders, immune deficiency diseases

CC (AIDS), mood disorders, respiratory disorders, arthritis, asthma,  
 CC haematopoietic disorders, neural disorders, skeletal disorders,  
 CC osteoporosis, metabolic disorders, cardiovascular disorders, endocrine  
 CC disorders or gastrointestinal disorders. The polypeptides are also useful  
 CC for identifying their binding partners. The polypeptide sequence represents a  
 CC gene encoding a human secreted protein (see descriptor line for gene  
 CC number and clone identification).  
 XX  
 SQ Sequence 1180 BP; 258 A; 304 C; 363 G; 250 T; 5 other:  
 Alignment Scores:  
 Pred. No.: 6.14e-94 Length: 1180  
 Score: 848.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0  
 US-10-082-502-20 (1-162) x AAX22112 (1-1180)  
 OY 1 ArgArgSerGluAspLeuHisCysGluAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 DB 573 CGGAGAGCCAGAGATCTCCACTGTGGAGCATGCAGGGCTCTGTGATGAACTAGACATGG 632  
 OY 21 GluIleAlaGluValAspProLysLysThrIleGluMetGlySerPheArgIleAsnPro 40  
 DB 633 GAAATGTCCAGAGTGGAGCCCAAGAAAGACCATTCAGATGGATCTTCCGATCAATCCA 692  
 OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 DB 693 GATGCGAGCCAGTCAGTGTGGAGGTGCTTATGCCCTCAGAGGCCACCATCCACAGAG 752  
 OY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 DB 753 CTGCTGGAGAGATATATGATCACCAGATGAAGAGACTATGGGAAACAGATTATCTTCACC 812  
 OY 81 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerGluLeuAspLeu 100  
 DB 813 CATCCGAAGAACTACGTACGTAGTGAGGCCGGAATGGAAATCCAGTCACTGACCTA 872  
 OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 DB 873 CAAGGCATCCGAATCGACATCATATATACCGGACCCCTCAAGTTCCGTGAGACCATT 932  
 OY 121 VALGLUGLUTYRGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValIys 140  
 DB 933 GTGGAGGAAATACGAGATCACTGTAATCTTTCCCGAGAGCTGCACATGTATAA 992  
 OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 DB 993 GACAAACTTTGGCAGTAAAGCAAGATCTTTGTGACACATGCCCTGCACATTCGACATGAT 1052  
 OY 161 GluLeu 162  
 DB 1053 GAGCTA 1058  
 RESULT 9  
 AAS21330  
 ID AAS21330 standard; cDNA; 1210 BP.  
 AC AAS21330;  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human cDNA sequence encoding for PRO4426 polypeptide.  
 XX  
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
 KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;  
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
 KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX

PN HQ200140466-A2.  
 XX  
 PD 07-JUN-2001.  
 XX  
 PF 01-DEC-2000; 2000MO-US32678.  
 XX  
 PR 01-DEC-1999; 99MO-US28301.  
 PR 01-DEC-1999; 99MO-US28634.  
 PR 02-DEC-1999; 99MO-US28551.  
 PR 02-DEC-1999; 99MO-US28554.  
 PR 02-DEC-1999; 99MO-US28565.  
 PR 09-DEC-1999; 99MO-US170262.  
 PR 16-DEC-1999; 99MO-US30095.  
 PR 20-DEC-1999; 99MO-US30911.  
 PR 20-DEC-1999; 99MO-US30999.  
 PR 30-DEC-1999; 99MO-US31243.  
 PR 06-JAN-2000; 2000MO-US00277.  
 PR 06-JAN-2000; 2000MO-US00376.  
 PR 11-FEB-2000; 2000MO-US03565.  
 PR 18-FEB-2000; 2000MO-US04341.  
 PR 18-FEB-2000; 2000MO-US04342.  
 PR 22-FEB-2000; 2000MO-US04414.  
 PR 24-FEB-2000; 2000MO-US04914.  
 PR 24-FEB-2000; 2000MO-US05004.  
 PR 01-MAR-2000; 2000MO-US05601.  
 PR 20-MAR-2000; 2000MO-US07377.  
 PR 21-MAR-2000; 2000MO-US07532.  
 PR 30-MAR-2000; 2000MO-US08439.  
 PR 17-MAY-2000; 2000MO-US13705.  
 PR 23-MAY-2000; 2000MO-US14042.  
 PR 30-MAY-2000; 2000MO-US14941.  
 PR 02-JUN-2000; 2000MO-US15264.  
 PR 10-NOV-2000; 2000MO-US30873.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Baker KP, Bersani M, DeForge L, Desnoyers L, Fillvaroff E, Gao W,  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 XX  
 DR WPI: 2001-408281/43.  
 DR P-PSDB; AAU12258.  
 XX  
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing  
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical  
 XX  
 PS Claim 3; Fig 173; 813pp; English.  
 XX  
 CC AAS21244-AAS21518 encode for novel human secretory and transmembrane  
 CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, to modulate biological activities of cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from  
 CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC of T-lymphocytes, the release of a cytokine from peripheral blood  
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding  
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy.  
 XX  
 XX Sequence 1210 BP; 261 A; 316 C; 379 G; 254 T; 0 other;

Alignment Scores:  
 Pred. No.: 6.35e-94 Length: 1210  
 Score: 848.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0  
 US-10-082-502-20 (1-162) x AAS21330 (1-1210)  
 QY 1 ArgArgSerGlnAspPleuHISCSYGLVALAcysrTgAlaLeuValAspGluLeuGluTrp 20  
 |||||||  
 DB 603 CGGAGGAGCCAGGATCTCCACTGTGGAGCATGCAGGGCTCTGGTGGATGAACTGGAATGG 662  
 QY 21 GluIleAlaGluValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
 |||||||  
 DB 663 GAATTTGCCAGGTGGAGCCCAAGAGACATTCAGATGGATCTTTCCGATCAATCCA 722  
 QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 |||||||  
 DB 723 GATGGCAGCCAGTCAGTGTGGAGGTGCTTATGCCCTCAGAGGCCACCTCAGAGAG 782  
 QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGluIleAspProSerThr 80  
 |||||||  
 DB 783 CTGCTGGAGGAGATATGTGACCGCATGAGAGATATGGGAAACGATTCCTTCCACC 842  
 QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerSerGluLeuAspLeu 100  
 |||||||  
 DB 843 CATCGCAAGAACTACGATACGTAGTGGGCCGGAATGAGAACTCAGTCACTGGACCTA 902  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 |||||||  
 DB 903 CAAGCATCCCAATCGACTCAGATATTAGCGCACCTCCAAAGTTGGTGTGAGAGCATTT 962  
 QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
 |||||||  
 DB 963 GTGAGGAAATACGAGGATGAACTCATTCATTTCCCGAGAGCTGACAAAGTTAA 1022  
 QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 |||||||  
 DB 1023 GACAAACTTTCAGTAAAGCGAAGATCTTGTGACCATGCTTCGACATATCCATGAT 1082  
 QY 161 GluLeu 162  
 |||||||  
 DB 1083 GAGCTA 1088  
 RESULT 10  
 AAX97884  
 ID AAX97884 standard; cDNA: 714 BP.  
 AC AAX97884;  
 XX  
 DT 23-SEP-1999 (first entry)  
 XX  
 DE Human secreted protein encoding cDNA #72.  
 XX  
 KW Secreted protein; human; cytostatic; thrombotic; osteopathic; forensic;  
 KW diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
 OS Homo sapiens.  
 XX  
 PN W09925825-A2.  
 PD 27-MAY-1999.  
 XX  
 PF 13-NOV-1998; 98MO-IB01862.  
 XX  
 PR 04-SEP-1998; 98US-0099273.  
 PR 13-NOV-1997; 97US-0066677.  
 PR 17-DEC-1997; 97US-0069957.  
 PR 09-FEB-1998; 98US-0074121.  
 PR 13-APR-1998; 98US-0081563.  
 PR 10-AUG-1998; 98US-0096116.



XX (GEST ) GENSET.  
 XX PA Bouqueloret L, Duclert A, Dumas Mline Edwards J;  
 XX PI WPI: 1999-347472/29.  
 XX DR P-PSDB: AAY36200.  
 XX DR P-PSDB: AAY36200.  
 XX PT Extended cDNAs encoding secreted proteins  
 XX PS Claim 1; Page 272; 307pp; English.  
 XX CC AA97913-x97906 represent extended cDNA's which encode novel human  
 CC secreted proteins (see AAY36129-136222) and which have cytosolic,  
 CC thrombotic and osteopathic activity. The extended cDNAs can be used to  
 CC express secreted proteins or parts of them or to obtain antibodies  
 CC capable of binding to the secreted proteins. They may also be used in  
 CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC Uses also include design of expression vectors and secretion vectors.  
 XX  
 SQ Sequence 714 BP; 203 A; 158 C; 202 G; 151 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2 26e-93 Length: 714  
 Score: 841.00 Matches: 161  
 Percent Similarity: 99.38% Conservative: 0  
 Best Local Similarity: 99.38% Mismatches: 1  
 Query Match: 99.17% Indels: 0  
 DB: Gaps: 0  
 US-10-082-502-20 (1-162) x AA97884 (1-714)  
 QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspLeuGluTrp 20  
 Db 93 CGAGAGACGACGAGATCCACTGTGAGACATGAGGCTGTGGATGACATGAGATG 152  
 QY 21 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
 Db 153 GAAATTTGCCAGCTGAGCCCAAGAACCATTCAGATGGATCTTCCGATCAATCCA 212  
 QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 Db 213 GATGCGACCGACTGAGTGTGAGAGTGTGCTTATGCCCGCTGAGAGCCACCTCCACAG 272  
 QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 Db 273 CTGCTGGAGAGATATGTCGACCGAGATGAGAGATGAGGAAACAGATTGATCTCCAC 332  
 QY 81 HisArgLysAsnTyrValArgValAlaGlyArgGlnGlyLysSerSerGluLeuAspLeu 100  
 Db 333 CATCCCAAGAACTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 392  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuValPheAlaCysGluSerIle 120  
 Db 393 CAAAGCATCCGAAATGATGATATTTACCGGACCCCTCAATTTGCGTGTGGAGCAT 452  
 QY 121 ValGluIleTyrGluAspLeuLeuIleGluPheSerArgGlnAlaAspAsnValLys 140  
 Db 453 GTGAGAGAAATACGAGATCAACTCATTTGAAATCTTTCCCGAGAGCTGACAAATGTTAA 512  
 QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 Db 513 GACAAACCTTTCAGTAAAGGACAGATCTTGTGACCATGCGCCCTGCATATGATGAT 572  
 QY 161 GluLeu 162  
 Db 573 GAGCTA 578  
 RESULT 11  
 ID AAX06970  
 XX AAX06970 standard; cDNA; 1069 BP.  
 AC AAX06970;

XX 10-MAY-1999 (first entry)  
 XX Mouse secretory peptide-9 (Zs1g9) orthologue cDNA.  
 DE Secretory peptide-9; Zs1g9; orthologue; mouse; tumour marker;  
 XX cancer; therapy; diagnosis; growth enhancer; ss.  
 KM Mus sp.  
 OS Key  
 XX CDS  
 FH Location/Qualifiers  
 FT 358..906  
 FT /\*tag= a  
 FT sig\_peptide 358..417  
 FT /\*tag= b  
 FT mat\_peptide 418..903  
 FT /\*tag= c  
 XX W09901554-A1.  
 XX  
 XX 14-JAN-1999.  
 PD  
 XX  
 XX 02-JUL-1998; 98WO-US13859.  
 XX  
 XX 17-JUN-1998; 98US-0099005.  
 PR 03-JUL-1997; 97US-0051704.  
 PR 03-JUL-1997; 97US-0088088.  
 PR 19-MAY-1998; 98US-0081338.  
 PR 19-MAY-1998; 98US-0085983.  
 PR 17-JUN-1998; 98US-0088899.  
 XX  
 XX (ZYMO ) ZYMOGENETICS INC.  
 XX Jaspers SR, Jelinek LJ, Shepard PO, Whitmore TE;  
 XX WPI: 1999-106055/09.  
 DR P-PSDB: AAM88476.  
 XX New mammalian secretory peptide-9 (Zs1g9) - used as a growth  
 PT enhancer for placenta, liver and heart, and as an indicator of  
 PT cancer  
 XX  
 XX Claim 2; Page 75-77; 85pp; English.  
 PS  
 XX This cDNA clone encodes novel mouse secretory peptide-9, or Zs1g9  
 CC (see AAM88476), an orthologue of novel human Zs1g9 (see AAM88465).  
 CC Human Zs1g9 is overexpressed in a number of tumours including  
 CC brain, liver, lung, oesophageal, stomach, colon, rectal, thyroid  
 CC and lymphoma tumors, and thus can be used as an indicator for  
 CC cancer. The invention provides polynucleotides (see AAX06968-70)  
 CC encoding Zs1g9 polypeptides (see AAM88469-77) including mature  
 CC polypeptides, other processed forms, variants and the mouse  
 CC orthologues. The Zs1g9 gene, or probes derived from it, can be  
 CC used to determine if Zs1g9 is present on chromosome 10, and if a  
 CC mutation has occurred. Antibodies raised against Zs1g9 can be  
 CC used as diagnostic agents to determine the presence of Zs1g9, and  
 CC thus the presence of cancer. They can also be labelled with  
 CC radioisotopes or fused with toxins and used to treat tumors  
 CC which overexpress Zs1g9. Antisense nucleotides derived from Zs1g9  
 CC cDNA can also be used to inhibit the growth of tumor cells.  
 CC Zs1g9 proteins can be used to enhance the growth or development of the  
 CC placenta, heart or liver.  
 XX  
 SQ Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 8 56e-92 Length: 1069  
 Score: 830.00 Matches: 158  
 Percent Similarity: 98.77% Conservative: 2  
 Best Local Similarity: 97.53% Mismatches: 2  
 Query Match: 97.88% Indels: 0  
 DB: Gaps: 0

US-10-082-502-20 (1-162) x AA06970 (1-1069)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 |||||  
 DB 418 CGAAGGAGCCAAAGATGTACACTGTGAGCTTGCGAGGCTCTGGTGGATGAAATGAGATCGG 477

QY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetCylSerPheArgIleAsnPro 40  
 |||||  
 DB 478 GAATATCCCGCGCGTGGAGCCCAAGAGACCATTCAGATGGATCTTCCAAATCAATCA 537

QY 41 AspGlySerGlnSerValIleGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 |||||  
 DB 538 GATGGCAGCCAGTCACTGTTGGAGAGTACCTTATGCCCCGCTCAGAGGCCACCTCAGAG 597

QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 |||||  
 DB 598 TTGCTTGAGAGAGTGTCTGACCGAATGAAAGAGAGTACGGGAAACGATTGACCTTCTAC 657

QY 81 HisArgLysAsnTyrValAlaGlyValGlyArgAsnGlyLysSerSerGluLeuAspLeu 100  
 |||||  
 DB 658 CACCGCAAGACATCACTACCGCTGAGCCCGGAATGGAATCCAGTGAACCTGACTTGA 717

QY 101 GlnGlyIleArgTyrIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 |||||  
 DB 718 CAGGGCATCCGAATTCAGATGATATCAGCGGACCTCTCAAGTTGGGTGGAGAGCAT 777

QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
 |||||  
 DB 778 GTGGAAGAAATACGAGATGAGCTTATCGAATTTCTTCCAGAGAGCGTGAACAGTTAAA 837

QY 141 AspLysLeuCysserLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 |||||  
 DB 838 GACAACTTTCAGATGAGCGGACAGATCTATGTGACCATCCCTCGACAGATCTCACAGAT 897

QY 161 GluLeu 162  
 |||||  
 DB 898 GAGCTA 903

RESULT 12  
 AA08294  
 ID AA08294 standard; DNA; 1069 BP.

XX AA08294:  
 AC  
 XX  
 DT 07-FEB-2000 (first entry)

XX Mouse ortholog gene encoding zsig9 secretory protein.  
 DE  
 XX  
 KW Secretory protein-9; mouse zsig9; ortholog; overexpression; antagonist;  
 KW antibody; antisense nucleotide; tumour; treatment; receptor;  
 KW radio-label; polypeptide toxin; down-regulation; diagnostic;  
 KW therapeutic; probe; cancer; brain; liver; detection; stomach;  
 KW lymphoma; ds.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key  
 FT CDS  
 FT Location/Qualifiers  
 FT 358..906  
 FT /\*tag- a  
 FT /product- "Mouse zsig9 secretory protein"  
 FT /note- "Overexpressed in tumours"  
 FT sig-peptide  
 FT 358..417  
 FT /\*tag- b  
 FT mat-peptide  
 FT 418..903  
 FT /\*tag- c  
 FT /label- Mature\_zsig9\_protein\_variant-4  
 PN MO9960405-A1.  
 XX  
 PD 25-NOV-1999.  
 XX  
 PF 19-MAY-1999; 99MO-US11107.  
 XX

PR 19-MAY-1998; 9805-0081183.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Moore EE, Taft DW;  
 XX  
 DR WPI: 2000-039447/03.  
 DR P-PSDB: AAY15136.  
 XX  
 PT Detecting tumors using antibodies, antagonists and antisense  
 PT nucleotides to secretory protein-9 (Zsig9) -  
 PS  
 PS Disclosure; Page 37-38; 45pp; English.  
 XX  
 CC The present DNA sequence is an ortholog encoding the secretory protein-9,  
 CC zsig9 derived from mouse. It is overexpressed in tumours. Antagonists,  
 CC antibodies and antisense nucleotides to zsig9 are useful for detecting  
 CC and treating tumours. The antagonist may be an antibody or receptor to  
 CC zsig9 and it may be radio-labelled or fused to a polypeptide toxin. It  
 CC can be used for down regulating the overexpression of zsig9. The gene  
 CC sequence can be used as nucleic acid probes to detect RNA encoding zsig9.  
 CC The zsig9 sequence facilitates improved diagnostic and therapeutic  
 CC techniques for detecting and treating cancers, especially of the brain,  
 CC liver, stomach, lymphoma etc., at an early stage.  
 CC  
 SQ Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other;

Alignment Scores:  
 Pred. No.: 8,56e-92 Length: 1069  
 Score: 830.00 Matches: 158  
 Percent Similarity: 98.77% Conservative: 2  
 Best Local Similarity: 97.53% Mismatches: 0  
 Query Match: 97.88% Indels: 0  
 DB: Gaps: 0

US-10-082-502-20 (1-162) x AA08294 (1-1069)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 |||||  
 DB 418 CGAAGGAGCCAAAGATGTACACTGTGAGCTTGCGAGGCTCTGGTGGATGAAATGAGATCGG 477

QY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetCylSerPheArgIleAsnPro 40  
 |||||  
 DB 478 GAATATCCCGCGCGTGGAGCCCAAGAGACCATTCAGATGGATCTTCCAAATCAATCA 537

QY 41 AspGlySerGlnSerValIleGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 |||||  
 DB 538 GATGGCAGCCAGTCACTGTTGGAGAGTACCTTATGCCCCGCTCAGAGGCCACCTCAGAG 597

QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 |||||  
 DB 598 TTGCTTGAGAGAGTGTCTGACCGAATGGAAGAGATGAGCGGAAACGATTCACCTTCTAC 657

QY 81 HisArgLysAsnTyrValAlaGlyValGlyArgAsnGlyLysSerSerGluLeuAspLeu 100  
 |||||  
 DB 658 CACCGCAAGACATCACTACCGCTGAGCCCGGAATGGAATCCAGTGAACCTGACTTGA 717

QY 101 GlnGlyIleArgTyrIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 |||||  
 DB 718 CAGGGCATCCGAATTCAGATGATATCAGCGGACCTCTCAAGTTGGGTGGAGAGCAT 777

QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
 |||||  
 DB 778 GTGGAAGAAATACGAGATGAGCTTATCGAATTTCTTCCAGAGAGCGTGAACAGTTAAA 837

QY 141 AspLysLeuCysserLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 |||||  
 DB 838 GACAACTTTCAGATGAGCGGACAGATCTATGTGACCATCCCTCGACAGATCTCACAGAT 897

QY 161 GluLeu 162  
 |||||  
 DB 898 GAGCTA 903

RESULT 13  
ABK36007  
ID ABK36007 standard; cDNA: 657 BP.  
XX  
XX  
AC ABK36007;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE cDNA sequence #398 encoding novel human secreted protein.  
XX  
KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;  
KW immune deficiency disorder; blood disorder; inflammatory disorder;  
KW infectious disorder; allergic condition; neurodegenerative disorder;  
KW liver fibrosis; coagulation disorder; gene therapy; antimicrobial;  
KW tumour; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN MO200177289-A2.  
XX  
PD 18-OCT-2001.  
XX  
PE 29-MAR-2001; 2001WO-US10232.  
XX  
PR 06-APR-2000; 2000US-195605P.  
XX  
PA (GENY ) GENETICS INST INC.  
XX  
PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
PI Merberg D, Treacy M, Agostino MJ, Bowman MR, Spaulding V, Wong GG;  
PI Clark HF, Fechtel K, Howes SH, Resnick RJ, Gulikova K, Graham JR;  
XX  
DR WPI: 2002-179322/23.  
XX  
XX Six hundred and twenty three polynucleotides derived from a variety of  
PT human tissue sources which encode secreted proteins, useful for  
PT treating immune deficiencies and disorders such as autoimmune disorders  
PT  
PS  
XX Claim 1; Page 297; 393pp; English.  
XX  
XX The present invention relates to the isolation of novel cDNA sequences  
CC which encode human secreted proteins. The cDNA sequences have been  
CC derived from a variety of human tissues. The invention also provides  
CC a method for producing proteins from these polynucleotide sequences.  
CC The proteins are useful for identifying compounds that modulate their  
CC activity and production. The sequences of the invention are  
CC useful for treating diseases such as hyperproliferative disorders  
CC (e.g. cancer), immune deficiency disorders (e.g. severe combined  
CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple  
CC sclerosis), blood disorders (e.g. thrombocytopenia), inflammatory  
CC disorders (e.g. arthritis), infectious disorders (e.g. hepatitis),  
CC allergic conditions (e.g. asthma), neurodegenerative disorders (e.g.  
CC Alzheimer's disease), liver fibrosis, coagulation disorders (e.g.  
CC haemophilia), and tumours. The polynucleotide sequences of the  
CC invention are also useful in gene therapy. ABK35610-ABK36232 represent  
CC the cDNA sequences of the invention that encode for novel human  
CC secreted proteins.  
XX  
SQ Sequence 657 BP; 177 A; 149 C; 185 G; 146 T; 0 other;

Alignment Scores:  
Pred. No.: 5,93e-92 Length: 657  
Score: 829.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.76% Indels: 0  
DB: 24 Gaps: 0

US-10-082-502-20 (1-162) x ABK36007 (1-657)  
OY 1 ArgArgSerGlnAspLeuHisCysGlyValAcysArgAlaLeuValAspGluLeuGluTrp 20  
|||||

DB 49 CGAGAGACCAAGATCTCCACTGTGGAGCATCAGAGGCTCTGTGATGACTAGAAATGG 108  
OY 21 GUILLEAGLInValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
|||||  
DB 109 GAAATTTGCCACAGCTGGACCCCAAGAAACATGACATGGATGATCTTCCGGATCATCA 168  
OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
|||||  
DB 169 GATTCACACCACTAGTGTGGAGGTGCGCTATGCCCCGTGACAGGCCACATCAGAG 228  
OY 61 LeuLeuGluGluIleLysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
|||||  
DB 229 CTGCTGGAGGAATATGTGACCGGATGAGAGGTATGGGAAACAGATTCATCTTCCACC 288  
OY 81 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerGluLeuAspLeu 100  
|||||  
DB 289 CATGCCAAGAACTACTAGTACTGTAGTGGCCCGAATGGAAATCCAGTAACTGGACCTA 348  
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
|||||  
DB 349 CAAGCCATCCGAAATGACATGACATATAGCCGACCCCTCAAGTTGCGTGGAGACAT 408  
OY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
|||||  
DB 409 GTGGAGGAATACAGAGATGAACTCATTTGAAATCTTTCCCGAGAGGCTACAAATGTAA 468  
OY 141 AspLysLeuCySerLysArgThrAspLeuCySerAspHisAlaLeuHisIleSerHisAsp 160  
|||||  
DB 469 GACAAACTTTCAGATGAGGAGAACAGATCTTTGTGACATGCCCTGCACATATCGCATGAT 528  
OY 161 GluLeu 162  
|||||  
DB 529 GAGCTA 534

RESULT 14  
ID AAX97837  
ID AAX97837 standard; cDNA: 832 BP.  
XX  
XX  
AC AAX97837;  
XX  
DT 23-SEP-1999 (first entry)  
XX  
XX Human secreted protein encoding cDNA #25.  
DE  
DE Secreted protein; human; cytostatic; thrombotic; osteopathic; forensic;  
KW diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9925825-A2.  
XX  
PD 27-MAY-1999.  
XX  
PE 13-NOV-1998; 98WO-IB01862.  
XX  
PR 04-SEP-1998; 98US-0099273.  
PR 13-NOV-1997; 97US-0066677.  
PR 17-DEC-1997; 97US-0069957.  
PR 09-FEB-1998; 98US-0074121.  
PR 13-APR-1998; 98US-0081563.  
PR 10-AUG-1998; 98US-0096116.  
XX  
PA (GENST ) GENSET.  
XX  
PI Bouguetleret L, Duclert A, Dumas Mline Edwards J;  
XX  
DR WPI: 1999-347472/29.  
XX  
DR P-PSDB: AAY36153.  
XX  
XX Extended cDNAs encoding secreted proteins  
XX  
XX Example 28; Page 194-195; 307pp; English.  
XX



BASE COUNT 723 a /note="cloning vector: pME18SFL13"  
ORIGIN 432 c 477 g 746 t

## Alignment Scores:

Pred. No.:	9,17e-16	Length:	2378
Score:	248.00	Matches:	50
Percent Similarity:	62.30%	Conservative:	26
Best Local Similarity:	40.98%	Mismatches:	44
Query Match:	29.28%	Indels:	2
DB:	9	Gaps:	1

US-10-082-502-21 (1-162) x AK094445 (1-2378)

```
QY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
    |||||
DB 80 GACGGACTGTACTACAGGGAGCAGATCCCTAGCTCAGTCGAGGCGTTCCTACCGAT 139
    ::::: |||||
QY 61 LeuLeuGluGluValAlaCysAspArgMetLysGluTyrGluGlnIleAspProSerThr 80
    |||||
DB 140 CTTTGGAGAAAGTCTGTGTGCGCAATGACGACTACAGCTGTGAGAGACCCGTGACG 199
    :::::
QY 81 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 100
    ::::: |||||
DB 200 AAGGAGAGAACTTTCAGAGATTCGCTCCTAGGAAAGAGACAAAATATACCAAGATTT 259
    ::::: |||||
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
    ::::: |||||
DB 260 AAAAATTTGATTTTATCTGATGCTTACAGACCTTGAATTTGCTGTGAAACTATA 319
    ::::: |||||
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
    ::::: |||||
DB 320 ATAGAAGATATGATGAAATATCTCATTATCGCCGAGGACACACTATCTAGCT 379
    ::::: |||||
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
    |||||
DB 380 GACAACTGTGCGAGTAAAAATCAGATCTGTGTGAATCTTCT-----GCTAATCAACT 433
    |||||
QY 161 GluLeu 162
    |||||
DB 434 GAGCTC 439
```

Search completed: March 31, 2003, 03:44:28  
Job time: 1677.7 secs



*	44607	46803: contig of 2197 bp in length
*	46804	46803: gap of unknown length
*	46904	49226: contig of 2323 bp in length
*	493227	49326: gap of unknown length
*	49327	51433: contig of 2107 bp in length
*	51434	51533: gap of unknown length
*	51534	53763: contig of 2230 bp in length
*	53764	53863: gap of unknown length
*	53864	56613: contig of 2750 bp in length
*	56614	56713: gap of unknown length
*	58773	58773: contig of 2060 bp in length
*	58774	58873: gap of unknown length
*	58874	61527: contig of 2654 bp in length
*	61528	61627: gap of unknown length
*	61628	63192: contig of 1565 bp in length
*	63193	63292: gap of unknown length
*	63293	65362: contig of 2070 bp in length
*	65363	65462: gap of unknown length
*	65463	65468: contig of 1976 bp in length
*	67439	67538: gap of unknown length
*	67539	70505: contig of 2967 bp in length
*	70506	70605: gap of unknown length
*	70606	74524: contig of 3919 bp in length
*	74525	74624: gap of unknown length
*	74625	76737: contig of 2113 bp in length
*	76738	76837: gap of unknown length
*	76838	80133: contig of 3296 bp in length
*	80134	80233: gap of unknown length
*	80234	83122: contig of 2889 bp in length
*	83123	83322: gap of unknown length
*	83323	86731: contig of 3509 bp in length
*	86732	86831: gap of unknown length
*	86833	89929: contig of 3098 bp in length
*	89930	90029: gap of unknown length
*	90030	93107: contig of 3078 bp in length
*	93108	93207: gap of unknown length
*	93208	97419: contig of 4212 bp in length
*	97420	97519: gap of unknown length
*	97520	101666: contig of 4147 bp in length

```

BASE COUNT      24650 a    22934 c    23569 g    23059 t    7454 others
ORIGIN

Alignment Scores:
Pred. No.:      3,42e-32      Length:      101666
Score:          429, 00      Matches:      97
Percent:        49.25%      Conservative: 1
Best Similarity: 48.74%      Mismatches:  2
Best Local Similarity: 50.65%      Indels:      99
Query Match:    2          Gaps:         1
OB:

```

US-10-082-502-21 (1-162) x AC109891 (1-101666)

QY	49	ValProYLrAlaArgSerGluIleHisLeuThrGluLeuGluGluValGlyAspArg	68
Db	97572	GTTCCTTATGCCCGCTCAGAGGCCCACTCAGCGAGTTCCTTGAGAGAGATATGTACCGCA	976313
QY	69	MetLysGluTYrGlyGluGlnIleLeuAspProSerThrHisArgLysAsnTYrValArgVal	88
Db	97632	ATGAAGGAGATNAGGGGAAACAGATTGACCCCTTACCACACCGGAAAGTACGTCCGTGT	97692
QY	89	ValSerAlaArgAsnGlyCysLeuSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp	108
Db	97692	GTCGGCCGCGAAGCGGAGATCCAGTCACTACAGACTTACAGGGGTATCCGATGCATTACAGAT	977513
QY	109	IleSerGlyThrLeuLeuPhe-----	115
Db	97732	ATCAATGGACACCTCAAGTTTCCGCTGACATTATGCGCTGCGGTGACATAGCTGGTTCTTG	978113

QY	115	-----	115
Db	97812	GCAGTTCATAGGAGATCTCTGGGCACAGATTGGAGACTCGGCGCGGATATGGTGTTCATG	97871
QY	115	-----	115
Db	97872	CTTTGTAACCCGGGACGTGGAGAGCAGCTGTTGGAGAATTGGCACTACTTCATACAGT	97931
QY	115	-----	115
Db	97932	TCTAGCGCAGCTGGAGCGATATCATATAAACCTCTCTATAAATCCCACTGCCACCAA	97991
QY	115	-----	115
Db	97992	AAACCGAGAGAGACTGAGCGGTGTCTGTGCCCCGGCTGAGATTCCTGGGTCACTCT	98051
QY	116	-----	116
Db	98052	CTGCTCTTGTGTCGCGCATGTGAACGATCTGTGGAGAGATPACGAGGATGAACCTTATGA	98111
QY	130	-----	130
Db	98112	ATTCTTTTCAGAGAGCGCTGACACAGTTTAAGACAACTTTGCACATAAGCGGACA	98166

RESULT 13	AF186113	649 bp	mRNA	linear	PRI 13-JAN-2000
LOCUS	AF186113				
DEFINITION	Homo sapiens putative secreted protein ZS1G9 (ZS1G9) mRNA, complete cds.				
ACCESSION	AF186113				
VERSION	AF186113.1	GI:6014631			
KEYWORDS					
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				

BASE COUNT	175 a	150 c	171 g	149 t	4 others
ORIGIN					
Alignment Scores:					
Pred. No.:	5.63e-17	Length:	649		
Score:	253.00	Matches:	48		
Percent Similarity:	100.00%	Conservative:	1		
Best Local Similarity:	97.96%	Mismatches:	0		

DEFINITION Rattus norvegicus clone CH230-319E6, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 47 unordered pieces.  
ACCESSION AC109891  
VERSION AC109891.3 GI:21738196  
KEYWORDS HTG: HTGS-PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 101666)  
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., All-oman,F.R., Allen,C.,  
Alebrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barbata,J., Benton,J., Blmage,K., Blankenburg,K., Bonin,D.,  
Bouch,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Butney,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dethorne,S.R., David,R.,  
Davis,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Doutharte,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Geo,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gottell,D.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,R., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Honsi,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Licharge,O., Liu,C., Liu,J., Liu,W., Loulseged,H.,  
Lozad,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Mageshwar,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokweto,S., Ogun,M., Okunodu,G.,  
Oregunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pull,L., Quiles,M., Ren,Y.,  
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,  
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Swalek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Umanan,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleceyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
TITLE Unpublished  
JOURNAL Direct Submission  
AUTHORS 2 (bases 1 to 101666)  
TITLE Worley,K.C.  
JOURNAL Direct Submission  
AUTHORS Submitted (08-FEB-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
REFERENCE 3 (bases 1 to 101666)  
AUTHORS Worley,K.C.  
JOURNAL Direct Submission  
AUTHORS Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
COMMENT On Jul 12, 2002 this sequence version replaced gi:18860222.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GQKt

Center clone name: CH230-319E6  
----- Summary Statistics  
Sequencing Vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 60229 bases at least Q40  
Consensus quality: 63959 bases at least Q30  
Consensus quality: 66730 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 47 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1  
\* 1423: contig of 1423 bp in length  
\* 1424  
\* 1523: gap of unknown length  
\* 1524  
\* 2972: contig of 1449 bp in length  
\* 2973  
\* 3072: gap of unknown length  
\* 3073  
\* 4717: contig of 1645 bp in length  
\* 4718  
\* 4817: gap of unknown length  
\* 4818  
\* 6632: contig of 1815 bp in length  
\* 6633  
\* 6732: gap of unknown length  
\* 6733  
\* 8040: contig of 1308 bp in length  
\* 8041  
\* 8140: gap of unknown length  
\* 8141  
\* 10043: contig of 1903 bp in length  
\* 10044  
\* 10143: gap of unknown length  
\* 10144  
\* 11237: contig of 1094 bp in length  
\* 11238  
\* 11337: gap of unknown length  
\* 11338  
\* 12362: contig of 1025 bp in length  
\* 12363  
\* 12463: gap of unknown length  
\* 12464  
\* 14100: contig of 1638 bp in length  
\* 14101  
\* 14200: gap of unknown length  
\* 14201  
\* 15840: contig of 1640 bp in length  
\* 15841  
\* 15940: gap of unknown length  
\* 15941  
\* 17053: contig of 1113 bp in length  
\* 17054  
\* 17153: gap of unknown length  
\* 17154  
\* 18836: contig of 1683 bp in length  
\* 18837  
\* 18936: gap of unknown length  
\* 18937  
\* 20137: contig of 1201 bp in length  
\* 20138  
\* 20237: gap of unknown length  
\* 20239  
\* 21736: contig of 1499 bp in length  
\* 21737  
\* 21836: gap of unknown length  
\* 21837  
\* 23619: contig of 1783 bp in length  
\* 23620  
\* 23719: gap of unknown length  
\* 23720  
\* 25907: gap of unknown length  
\* 25908  
\* 26007: contig of 2188 bp in length  
\* 25908  
\* 26007: gap of unknown length  
\* 26008  
\* 27406: contig of 1399 bp in length  
\* 27407  
\* 27506: gap of unknown length  
\* 27507  
\* 28837: contig of 1331 bp in length  
\* 28838  
\* 28937: gap of unknown length  
\* 28938  
\* 30229: contig of 1292 bp in length  
\* 30229  
\* 30329: gap of unknown length  
\* 30330  
\* 32012: contig of 1683 bp in length  
\* 32013  
\* 32112: gap of unknown length  
\* 32113  
\* 33198: contig of 1086 bp in length  
\* 33199  
\* 33298: gap of unknown length  
\* 33299  
\* 35519: contig of 2221 bp in length  
\* 35520  
\* 35620: gap of unknown length  
\* 35621  
\* 37677: contig of 2058 bp in length  
\* 37678  
\* 37777: gap of unknown length  
\* 37778  
\* 38935: contig of 1159 bp in length  
\* 38936  
\* 39035: gap of unknown length  
\* 39037  
\* 40273: contig of 1237 bp in length  
\* 40274  
\* 40373: gap of unknown length  
\* 40374  
\* 41906: contig of 1533 bp in length  
\* 41907  
\* 42006: gap of unknown length  
\* 42007  
\* 44506: contig of 2500 bp in length  
\* 44507  
\* 44506: gap of unknown length





Oy 116 ----- 116  
 Db 171727 AAGTGAAGCAAGAGCCTTATATGCTGTCACCCATTCTCCCTGGGTTGGCAG 171668  
 Oy 117 CysGUSerIleValGluGluTyrGluAspGluLeuIleGluPheSerArgLysAla 136  
 Db 171667 TGTGAGACATTTGTGAGAAATACGAGATGAACATTTGAATCTTTCCGAGAGGCT 171608  
 Oy 137 AspaNValIysAspIysLeuCySerIysArgThr 148  
 Db 171607 GACATGTTTAAAGACAAACTTTGCAGTAAGCCACACA 171572  
 RESULT 10  
 AC023500  
 LOCUS Homo sapiens chromosome 12 clone RP11-183H16, WORKING DRAFT  
 DEFINITION  
 AC023500 184762 bp DNA linear HTG 24-AUG-2002  
 AC023500  
 VERSION AC023500.32 GI:22095070  
 KEYWORDS  
 SOURCE HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 ORGANISM human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 184762)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oshan,F.R., Allen,C.,  
 Albrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T.,  
 Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,  
 Bouck,J., Bowle,S., Brieve,M., Brown,E., Brown,M., Bryant,N.P.,  
 Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Carrott,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Kocha,S., Durbin,K.J.,  
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabis,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
 Homsl,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Jackson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kravtovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,  
 Lorado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maneswari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,  
 Miner,G., Miner,Z., Mitchell,J., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokenwo,S., Ogub,M., Okunolu,G.,  
 Orangunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peter,L., Pichens,R., Primus,E., Pu,L.L., Qulles,M., Ren,Y.,  
 Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Rulz,S., Saverly,G.,  
 Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,  
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Sytek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Usmani,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,  
 Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstock,G. and Gibbs,R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 184762)  
 TITLE  
 AUTHORS Worley,K.C.  
 JOURNAL Direct Submission  
 Submitted (15-FEB-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One

REFERENCE  
 3 (bases 1 to 184762)  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Baylor Plaza, Houston, TX 77030, USA  
 Worley,K.C.  
 Direct Submission  
 Submitted (24-AUG-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Aug 4, 2002 this sequence version replaced gi:22094227.  
 -----  
 Center: Baylor College of Medicine  
 Genome Center  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Drafting Center Code: BCM  
 Contact: hgsc-help@bcm.tmc.edu  
 -----  
 Project Information  
 Center project name: HAF  
 Center clone name: RP11-183H16  
 -----  
 Summary Statistics  
 Sequencing vector: Plasmid;  
 Sequencing vector: M13;  
 Chemistry: Dye-primer Bodipy: 2% of reads  
 Chemistry: Dye-terminator Big Dye: 98% of reads  
 Assembly program: Phrap: version 0.990329  
 Consensus quality: 183845 bases at least Q40  
 Consensus quality: 184849 bases at least Q30  
 Consensus quality: 185447 bases at least Q20  
 Estimated insert size: 159026; sum-of-contigs estimation  
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
 \* NOTE: This is a "working draft" sequence. It currently  
 \* consists of 9 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1 2040: contig of 2040 bp in length  
 \* 2141 2140: contig of 2030 bp in length  
 \* 4171 4270: contig of 2030 bp in length  
 \* 4271 4270: gap of unknown length  
 \* 8295 8294: contig of 4024 bp in length  
 \* 8395 8394: gap of unknown length  
 \* 14197 14196: contig of 5802 bp in length  
 \* 14297 14296: gap of unknown length  
 \* 14297 14296: contig of 18780 bp in length  
 \* 33077 33176: gap of unknown length  
 \* 33177 33176: contig of 19119 bp in length  
 \* 52295 52295: gap of unknown length  
 \* 52296 52295: gap of unknown length  
 \* 84185 84185: contig of 31790 bp in length  
 \* 84186 84185: gap of unknown length  
 \* 84286 120023: contig of 35738 bp in length  
 \* 120024 120123: gap of unknown length  
 \* 120124 184762: contig of 64639 bp in length.  
 \*  
 \* Location/Qualifiers  
 \* 1. 184762  
 \* /organism="Homo sapiens"  
 \* /db\_xref="taxon:9606"  
 \* /chromosome="12"  
 \* /clone="RP11-183H16"  
 \* /location="12p11.23"

## ALIGNMENT SCORES:

Pred. No.: 2.96e-36 Length: 184762  
 Score: 472.00 Matches: 98  
 Percent Similarity: 65.13% Conservative: 1  
 Best Local Similarity: 64.47% Mismatches: 1  
 Query Match: 55.73% Indels: 52  
 DB: 2 Gaps: 1

RESULT 9  
AC073896/c

LOCUS AC073896 182350 bp DNA linear HTG 24-AUG-2002

DEFINITION Homo sapiens chromosome 12 clone RP11-977G19, WORKING DRAFT

ACCESSION AC073896

VERSION AC073896.37 GI:22450359

KEYWORDS HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 182350)

AUTHORS Munzy,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C., Alabrooks,S.L., Amaralunge,H.C., Are,J.R., Ayela,M., Banks,T., Barbara,J., Benton,J., Blmage,K., Blankenburg,K., Bonnin,D., Bouck,J., Boyle,S., Bileva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Dellaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durdin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frintz,P., Gabisl,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsl,F.P., Howard,S., Huber,J., Hultyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Kallason,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kretovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Ilean,C., Liu,J., Liu,W., Louiseged,H., Locardo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., McWhiney,E., McLeod,M.P., Mendon,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokhenkwo,S., Oguh,M., Okwuonu,G., Oreguene,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojnbokan,I., Roife,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonalike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabot,P., Tameisa,A., Tameisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telirod,B., Thomas,N., Thomas,S., Umanji,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Wolley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gdbps,R.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 182350)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (01-JUL-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 182350)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (24-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Aug 23, 2002 this sequence version replaced gi:2296903.

----- Genome Center -----

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Drafting Center Code: BCM

```

Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: HBNI
Center clone name: RP11-977G19
----- Summary Statistics -----
Sequencing vector: Plasmid;
Sequencing vector: M13;
Chemistry: Dye-primer Bodipy: 15% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 179279 bases at least Q40
Consensus quality: 180581 bases at least Q30
Consensus quality: 181073 bases at least Q20
Estimated insert size: 174258: sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This is a "working draft" sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
      1          2893: config of 2893 bp in length
      *          2894          gap of unknown length
      *          2984          config of 3693 bp in length
      *          6686          config of 10465 bp in length
      *          6687          gap of unknown length
      *          6787          config of 10465 bp in length
      *          17251         config of 10465 bp in length
      *          17352         gap of unknown length
      *          17352         config of 68543 bp in length
      *          85895         gap of unknown length
      *          85895         config of 96356 bp in length.
      *          Location/Qualifiers
            1..182350
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /chromosome="12"
              /clone="RP11-977G19"
BASE COUNT   46453 a 45416 c 43855 g 46174 t    452 others
ORIGIN
Alignment Scores:
Pred. No.:           2,92e-36                      Length:          182350
Score:             472.00                        Matches:             98
Percent Similarity: 65.13%                       Conservative:       1
Best local Similarity: 64.47%                     Mismatches:        1
Query Match:      55.73%                          Indels:             52
DB:               2                                Gaps:              1
US-10-082-502-21 (1-162) x AC073896 (1-182350)
OY 49 ValProTyrAlaAspSerGluAlaHisIleThrGluLeuLeuGluGluValCysAspArg 68
|||||
Db 172027 GTGCCTATAGCCCCCTCGAAGGCCACCTCAACAAGGTGGTGGAAGAATGTGACC GG 171968
OY 69 MetLysGluTyrGlyGlnGlnIleAspProSerThrHisArgLysAsnTyrValAlaGval 88
|||||
Db 171967 ATGAAGGAGTATGGGGAACAGATTGATTCCTCCACCACCATGCAAGAACTACGTACG 171908
OY 89 ValSerArgAnsiGluSerSerGluLeuAspLeuGlnIleArgIleAspSerAsp 108
|||
Db 171907 GTGGCCGGAGATGGAGATTCAGTAGACTGACCTACAAAGCAATCCGAATCGACTCAGAT 171848
OY 109 IleSerGlyThrIleuLysPheala
|||||
Db 171847 AATAAGCGCACACCTCAAGTTTGGCGGTAGCATGTAGCGAAATGGTAAGCTGTTCGACAT 171788
OY 116
Db 171787 TAAAGGGGTGTTGGAGAACCAATTAGAAATGTAAGGTGATATCTTGTCCCTCTTGG 171728

```

```

REFERENCE TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 155023)
AUTHORS Worley,K.C.
JOURNAL Direct Submission
TITLE Submitted (11-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 155023)
AUTHORS Worley,K.C.
JOURNAL Direct Submission
TITLE Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On May 3, 2002 this sequence version replaced gi:20335511.
-----
Center: Baylor College of Medicine
Genome Center
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: HALV
Center clone name: RP11-348M3
-----
Summary Statistics
Sequencing vector: M13
Chemistry: Dye-Primer Bodipy: 5% of reads
Chemistry: Dye-terminator Big Dye: 95% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 141522 bases at least Q40
Consensus quality: 147097 bases at least Q30
Consensus quality: 150442 bases at least Q20
Estimated insert size: 150706; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
*
* 1 3033: contig of 3033 bp in length
* 3034 3133: gap of unknown length
* 3134 5971: contig of 2838 bp in length
* 5972 6071: gap of unknown length
* 6072 10269: contig of 4198 bp in length
* 10270 10369: gap of unknown length
* 10370 13009: contig of 2640 bp in length
* 13010 13109: gap of unknown length
* 13110 17573: contig of 4464 bp in length
* 17574 17674 17673: gap of unknown length
* 17674 24650: contig of 6677 bp in length
* 24651 24750: gap of unknown length
* 24751 31778: contig of 7028 bp in length
* 31779 31878: gap of unknown length
* 31879 42263: contig of 10385 bp in length
* 42264 42363: gap of unknown length
* 42364 51104: contig of 8741 bp in length
* 51105 51204: gap of unknown length
* 51205 68306: contig of 17102 bp in length
* 68307 68406: gap of unknown length
* 68407 109481: contig of 41075 bp in length
* 109482 109581: gap of unknown length
* 109582 155023: contig of 45442 bp in length.
Location/Qualifiers
1. 155023
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-348M3"

```

BASE COUNT	40008 a	36207 c	35987 g	41693 t	1128 others					
ORIGIN										
Alignment Scores:										
Pred. No.:	1.32e-36	Length:	155023							
Score:	474.50	Matches:	111							
Percent Similarity:	39.02%	Conservative:	1							
Best Local Similarity:	38.68%	Mismatches:	2							
Query Match:	56.02%	Indels:	173							
DB:	2	Gaps:	2							
US-10-082-502-21 (1-162) x AC025574 (1-155023)										
OY 49	ValPcYrRAlAArgSerGluuAhIstLeuThrgLuLeuGluValcysAspArg	68								
Db 56120	GTGCGCTTATGCCGCTCAGAGGCCACCTCACAGAGCTGCTGAGAGATATGTGACCG	56061								
OY 69	MelTysGluTYrGluGluGlnLeaSPProSerThrHisArgLysAsnTyValArgVal	88								
Db 56060	ATGAAGAGATAGGGGAGACAGATATGATCTTCACCCATCGCAAGACTACGTACTGT	56001								
OY 89	ValSerArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyLeaArgLLeaSPSerAsp	108								
Db 56000	GTGGCCCGGAATGGAGAATCCGTAACGTGACCTACAAAGCATCCGATCGACTCAGAT	55941								
OY 109	IleSerGlyThrLeuLysPheIleA-----	116								
Db 55940	ATTACGGGACCCCTAAAGTTGGCTGAGCTATGGGAATCGGTACCTGGTCTTGACAT	55881								
OY 116	-----	116								
Db 55880	TAAGGGTTGTTGGAGAACCAATTAGAAAGTTAGGCTGATATCTGTTCCTCTTTGG	55821								
OY 116	-----	116								
Db 55820	AAGTGAAGGCAAGAGCCCTTATTTGCCCTGTGCACCCATTCTCCCTGGGTTGGCAG	55761								
OY 117	CysGluSerIleValGluGluTYrGluAspGluLeuIleGluPheSerArgGluA	136								
Db 55760	TGTGAGACATTTGTGAGGAATACAGAGATGAGACTCATTTGAATCTTTTCCGAGGGCT	55701								
OY 137	AspAsnValLysAspLysLeuCysSerLysArgThr-----	148								
Db 55700	GACATATGTTAAAGCAAACTTGGCAGTAGAGCAAGAGATAGCTGCCCACTTATATC	55641								
OY 148	-----	148								
Db 55640	CTGTTCTAGTAGCTCTGTGGAACCTGACATTCCTAACTCACCCATACACCCATCCCA	55581								
OY 148	-----	148								
Db 55580	GAAACCTCTGGATTAGGTCAAGAGACTGTCTCCATTTCCTTGGCTTAGGGGACTTGGG	55521								
OY 148	-----	148								
Db 55520	GTGTTCTCCACATTTTCTGTGACGCTTCAGAAATTTGTGCCCTTGACCTGAGATACA	55461								
OY 148	-----	148								
Db 55460	GGTGAGAGACAGTGGGAGATGTTTATCATTTATATAGCATATATTTCTCGAGATGTGCT	55401								
OY 148	-----	148								
Db 55400	TGGTGTGCTTTGTGTACAGCTTGTCTGGGAATTTTCTGTGATCTGCGCAGATACAGA	55341								
OY 149	-----	155								
Db 55340	TGAGATAGGCATATATCTTTAAATGTTTGCCATTTGCAGATCTTTGTGACCATCCCTG	55281								
OY 156	HisArgSerHisAspGluLeu 162									
Db 55280	CACATATGCAATGATGAGCTA 55260									

[illegible]

RESULT	7
AX464040	
LOCUS	AX464040
DEFINITION	Sequence 173 from Patent WO0140466.
ACCESSION	AX464040
VERSION	AX464040.1 GI:21899037
KEYWORDS	.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE
1	Baker, K. P., Beresini, M., DeForge, L., Denoyers, L., Flivaroff, E., Gao, M.O., Gerritsen, M.E., Goddard, A., Godowski, P. J., Guney, A. L., Sherwood, S., Smith, V., Stewart, T. A., Tunnas, D., Watanabe, C. K., Wood, M. L. and Zhang, Z.	Secreted and transmembrane polypeptides and nucleic acids encoding same	Patent: WO 01/40466-A 173 07-JUN-2001;		
	Genentech Inc. (US)				
	Location/Qualifiers				
	1. .1210				

Alignment Scores:	
Pred. No.:	1.75e-75
Score:	830.00
Percent Similarity:	98.77%
Best Local Similarity:	97.55%
Query Match:	97.99%
OB:	6
Length:	1210
Matches:	158
Conservative:	2
Mismatches:	2
Indels:	0
Gaps:	0

US-10-082-502-21 (1-162) x AX464040 (1-1210)

Qy	1	ArgdgsSergIAspLeuNIhCSyG61VALAcysArgIaIeUvaIAspGluLeuGIITp	20
Db	603	CGAGAGACCAGAGATCTCCATGTGTGGAGCATGCAAGGGCTCTGGTGGATGAACTGAAATGG	665
Qy	21	GIuIeIaIaArgVaIAspProLyIaYthrIleGIImetGlySerPheArgIleAspPro	40
Db	663	GAATTTGGCCAGGTGGAGCCCAAGAACCATTCAGATGGGATCTTTCCGGATCAATCA	722
Qy	41	AspGlySerGIInserVaIaIGluVaIProTyrIaIAspSergIuIaIhIAspLeuThGIu	60
Db	723	GATGGCACCCAGTCAGTGTGTGGAGGTGGCTTATTCGCCGCTCAGAGGCCCACTCACAGAG	783
Qy	61	LeuLeuGIuGIuVaIcysAspArgIaMetLyS6IuYtrGIuGIuGIuIleAspProSerThr	80
Db	783	CTGGCTGGAGAGATATGTGACCCGATGAAAGAGATATGGGGAAACAGATATGATCCCTTCCACC	842

OY	81	HISRGVYSANTRYVALACGVALVLSERARGNGLYGSERSCILUEUASPLEU	100
Db	843	CATGGCAGAACTRCAGTGTGTACTGTGGCCCGGAATGCAGAAATCCAGTAATTCGACCTA	902
OY	101	GINCILYLTLIARGLLEASPSSERASPILESERGILYTHRLIEULYSPHEALACSGLUSERIE	120
Db	903	CAAGCATTCGGAATCGCATCAGATTATTAGCGGACCCCTCAAGTTTCCGTGTAGAGACAAT	962
OY	121	VALIDLUDLYRGIUNASPGILUEULLEGILUPHESSETRARGIUALIASPSANVALYS	140
Db	963	GTTGGAGGAAATTCGAGGATGAACTCATTTGATTTTTCCCAGAGGCTGTACCAATGTATAA	1022
OY	141	ASPELYSLEUCYSSEIRLYSARGTHRASPLIEUCYSASPDLIALALEUHLISARGSERHISAP	160
Db	1023	GACAAACTTTTGCACTAAGCAGAACGATCTTTGTGGACCATTGCCCTGCACATATTCGCATGAT	1082
OY	161	GLIUEN 162	
Db	1083	GAGCTA 1088	

AC025574/c	AC025574	155023 bp	DNA	linear	HTG_09-MAY-2002
LOCUS					
DEFINITION	Homo sapiens chromosome 12 clone RP11-348M3, WORKING DRAFT				
ACCESSION	AC025574				
VERSION	AC025574.13 GI:20428723				
KEYWORDS	HTG; HTGS-PHASE1; HTGS-DRAFT.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				

REFERENCE  
AUTHORS  
Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.  
1 (bases 1 to 155023)  
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,

Babadi, J., Benton, J., Blmage, K., Blankenburg, K., Bonnin, D.,  
Bouck, J., Bowles, S., Briveau, M., Brown, E., Brown, M., Bryant, N.P.,  
Bunyah, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Davis, M.L., Davis, C., Davey-Carroll, L., Deckerich, D.A.,  
Delany, R.K., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
Donthwaite, K.J., Draper, H., Dugan-Roches, S., Durbin, K.J.,  
Eamhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Franz, P.,  
Gabals, A., Geo, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, R.,  
Harris, C., Harris, K., Hart, M., Havlik, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hoques, M., Holloway, C., Hollins, B.,  
Honsi, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolyet, S., Joudah, S.,  
Karlsson, E., Kelly, S., Khan, U., King, L., Korvan, J., Kover, C.,  
Kralovic, J., Kunesht, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Lounsged, H.,  
Lozardo, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Mageswarari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzker, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabdat, K., Morgan, M., Morris, S.,  
Mosier, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokweto, S., Ogulu, M., Okunodu, G.,  
Oregunye, N., Oviedo, R., Pace, A., Payton, B., Peary, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rives, M., Rojas, A., Rotjokan, I., Rolfe, M., Ruiz, S., Savery, G.,  
Schierer, S., Scott, G., Shen, H., Shoshtari, N., Slison, I.,  
Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Swalek, A., Tabori, P., Tamerias, A., Tamerias, K., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, O.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleceyk, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, Y., Zorilla, S., Nelson, D.,  
Winstock, G. and Gibbs, R.

REFERENCE 1  
AUTHORS Yokoyama-Kobayashi, M., Yamaguchi, T., Sekine, S. and Kato, S.  
TITLE Selection of cDNAs encoding putative type II membrane proteins on the cell surface from a human full-length cDNA bank  
JOURNAL Gene 228 (1-2), 161-167 (1999)  
MEDLINE 99173880  
REFERENCE 2 (bases 1 to 814)  
AUTHORS Kato, S.  
TITLE Direct Submission  
JOURNAL Submitted (16-JUN-1998) Seishi Kato, Research Institute of National Rehabilitation Center for the Disabled, Department of Rehabilitation Engineering, 4-1 Namiki, Tokorozawa, Saitama 359-8555, Japan (E-mail: seishi@rehab.go.jp, Tel:042-995-3100(ex.2568), Fax:042-995-3132)  
LOCATION/Qualifiers  
FEATURES  
SOURCE 1. 814  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="HP10390"  
/issue\_type="gastric adenocarcinoma"  
/clone\_id="PKA1-meta-1"  
145..693  
/codon\_start=1  
/product="type II membrane protein"  
/protein\_id="BAAF6498.1"  
/db\_xref="GI:4586840"  
/translation="MKMGWLLALLGALGTANARRSODLHCACRALVDELEWIAQ VDKPKTIOMGSFRINPDGSSQVVEVPAREAHITELLEICDKMKYGGQIDPSTHR KNYRVVGRNGSESLDGLGIRIDSDISGLFKACSEIVEYEDLEIFRSREADNVK DKICKRTDCLDHALHSHDEL"  
BASE COUNT 210 a 194 c 241 g 169 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.07e-75 Length: 814  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.99% Indels: 0  
Gaps: 0  
US-10-082-502-21 (1-162) x AB015631 (1-814)  
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
DB 205 CGAGAGACCGAGATCTCCACTGTGAGACAGGAGGCTGTGGATGAGTCAAGATGG 264  
QY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
DB 265 GAAATTCGCCAGGTGGAGCCCAAGAGACCATTCAGATGGGATCTTCCGGATCAATCCA 324  
QY 41 AspGlySerGlnSerValIleGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
DB 325 GATGGACGACGATGAGTGGTGGAGTGCCTTATGCCCCGCTCAGAGGCCCACTCAGAG 384  
QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
DB 385 CTGCTGGAGAGATATGATGACCGGATGAAGAGATATGGGGAACAGATTGATCTCCACC 444  
QY 81 HisArgLysAsnTyrValArgValIleSerArgAsnGlyGlySerGluLeuAspLeu 100  
DB 445 CATGGCAAGACTACGATCGTGTAGTGGGCGGAATGAGATCCAGTGAACCTGGACCTA 504  
QY 101 GlnGlyIleArgGlyLeaSerSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 505 CAAAGCATCCCAATCGACTCGATATTAAGCGGCACCTCAAGATTGGCTGTGAGAGCAAT 564  
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
DB 565 GTGAGAGAAATACGAGATGAACCTCATTTGATTTCCCGAGAGCGTGAACAATGTTAA 624  
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160

DB 625 GACAAACTTTCAGTAGACGAGACAGATCTTGTGACCATGCCCTGCACATATCGCATGAT 684  
QY 161 GluLeu 162  
DB 685 GAGCTA 690  
RESULT 6  
AY032624  
LOCUS AY032624 1080 bp mRNA linear PRI 18-APR-2002  
DEFINITION Homo sapiens saposin-like protein mRNA, complete cds.  
ACCESSION AY032624  
VERSION AY032624.1 GI:20196198  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1080)  
AUTHORS Bornhauser, B.C., Olsson, P.-A. and Lindholm, D.  
TITLE NSAP is a novel saposin-like protein that interacts with M1R and stimulates neurite outgrowth  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1080)  
AUTHORS Olsson, P.-A. and Lindholm, D.  
TITLE Direct Submission  
JOURNAL Submitted (13-APR-2001) Neuroscience, Uppsala University, Husargatan 3, Uppsala 75123, Sweden  
LOCATION/Qualifiers  
FEATURES  
SOURCE 1. 1080  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/map="12q15"  
419..967  
/note="NSAP; encodes a type 2 membrane signal and a putative ER retention motif"  
/codon\_start=1  
/product="saposin-like protein"  
/protein\_id="AAK8148.1"  
/db\_xref="GI:20196198"  
/translation="MKMGWLLALLGALGTANARRSODLHCACRALVDELEWIAQ VDKPKTIOMGSFRINPDGSSQVVEVPAREAHITELLEICDKMKYGGQIDPSTHR KNYRVVGRNGSESLDGLGIRIDSDISGLFKACSEIVEYEDLEIFRSREADNVK DKICKRTDCLDHALHSHDEL"  
BASE COUNT 238 a 269 c 343 g 230 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.52e-75 Length: 1080  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.99% Indels: 0  
Gaps: 0  
US-10-082-502-21 (1-162) x AY032624 (1-1080)  
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
DB 479 CGAGAGACCGAGATCTCCACTGTGAGACAGGAGGCTGTGGATGAGTCAAGATGG 538  
QY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
DB 539 GAAATTCGCCAGGTGGAGCCCAAGAGACCATTCAGATGGGATCTTCCGGATCAATCCA 598  
QY 41 AspGlySerGlnSerValIleGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
DB 599 GATGGACGACGATGAGTGGTGGAGTGCCTTATGCCCCGCTCAGAGGCCCACTCAGAG 658  
QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
DB 659 CTGCTGGAGAGATATGACCGGATGAAGAGATATGGGGAACAGATTGATCTCCACC 718

\* 150247 150346: gap of unknown length  
 \* 150347 153410: contig of 3064 bp in length  
 \* 153411 153510: gap of unknown length  
 \* 153511 155739: contig of 2229 bp in length  
 \* 155740 155839: gap of unknown length  
 \* 155840 158198: contig of 2359 bp in length.

FEATURES  
 source  
 1. 158198  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /clone="RP11-764L14"

BASE COUNT 46057 a 30922 c 31747 g 49030 t 442 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.39e-74 Length: 158198  
 Score: 847.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0

US-10-082-502-21 (1-162) x AC012013 (1-158198)

QY 1 ArgArgSerGlnAspLeuH1scysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 |||||  
 Db 152171 CGAAGAGCCAGCATCTACACCTGTGAGGCTGTGATGATGATGATGATGATG 152230  
 QY 21 GluIleAlaArgValAspProlySlyThrIleGlnMetGlySerPheArgIleAspPro 40  
 |||||  
 Db 152231 GAATATGCCCGGTGAGCCCAAGAACCATTCACATGGATGCTCCGATCAATCCA 152290  
 QY 41 AspGlySerGlnSerValAlaGluValProTyAlaArgSerGluAlaHisLeuThrGlu 60  
 |||||  
 Db 152291 GATGCAAGCCAGCATCTGTGAGGATCACTTATGCCCTCAGAGGCCACCTCACAGAG 152350  
 QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyGlyGluGlnIleAspProSerThr 80  
 |||||  
 Db 152351 TTGCTTGAGAGAGTGTGTGACCAATGAAGAGATGACGGGAGACAGATGACCTTCTAC 152410  
 QY 81 HisArgLysAsnTyValAlaArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 100  
 |||||  
 Db 152411 CACCGCAAGAACTACCTACGCGGTGAGCCGGAATGGAATCCGATGTAATCACTTA 152470  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 |||||  
 Db 152471 CAGGCAATCCGATTCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 152530  
 QY 121 ValGluGluTyGlyLysAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
 |||||  
 Db 152531 GTGCAAGAACTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 152590  
 QY 141 AspLysLeuLysSerLysArgThrAspLeuLysAspHisAlaLeuHisArgSerHisAsp 160  
 |||||  
 Db 152591 GACAAACTTTGAGTAAAGGAGACAGATCTATGTGACCATGCTCGACATCTCAGAT 152650  
 QY 161 GluLeu 162  
 |||||  
 Db 152651 GAGCTA 152656

RESULT 4  
 AX440456 814 bp DNA linear PAT 28-JUN-2002  
 LOCUS AX440456 Sequence 309 from Patent WO0190154.  
 DEFINITION  
 ACCESSION AX440456  
 VERSION AX440456.1 GI:21665266  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1

AUTHORS Xu, J., Mitcham, J. L., Harlocker, S. L., Dillon, D. C., Secrist, H.,  
 Lodes, M. J., Algate, P. A., Fling, S. P., Mannion, J., Benson, D. R. and  
 Carter, D.  
 TITLE Compositions and methods for the therapy and diagnosis of ovarian  
 cancer  
 JOURNAL Patent: WO 0190154-A 309 29-NOV-2001;  
 CORIXA CORPORATION (US)

FEATURES  
 source  
 1. 814  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

BASE COUNT 210 a 194 c 241 g 169 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.07e-75 Length: 814  
 Score: 830.00 Matches: 158  
 Percent Similarity: 98.77% Conservative: 2  
 Best Local Similarity: 97.53% Mismatches: 2  
 Query Match: 97.99% Indels: 0  
 DB: Gaps: 0

US-10-082-502-21 (1-162) x AX440456 (1-814)

QY 1 ArgArgSerGlnAspLeuH1scysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 |||||  
 Db 205 CGAGAGAGCCAGCATCTCCACCTGTGAGGCTGTGATGATGATGATGATGATGATG 264  
 QY 21 GluIleAlaArgValAspProlySlyThrIleGlnMetGlySerPheArgIleAspPro 40  
 |||||  
 Db 265 GAATATGCCCGGTGAGCCCAAGAACCATTCACATGGATGCTCCGATCAATCCA 324  
 QY 41 AspGlySerGlnSerValAlaGluValProTyAlaArgSerGluAlaHisLeuThrGlu 60  
 |||||  
 Db 325 GATGCAAGCCAGCATCTGTGAGGATCACTTATGCCCTCAGAGGCCACCTCACAGAG 384  
 QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyGlyGluGlnIleAspProSerThr 80  
 |||||  
 Db 385 CTGCTGAGAGAGATGATGTGACCGCATGAAGAGATGAGGAGACAGATGATCTTCAC 444  
 QY 81 HisArgLysAsnTyValAlaArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 100  
 |||||  
 Db 445 CATGCAAGAACTACCTACGCGGTGAGCCGGAATGGAATCCGATGTAATCACTTA 504  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 |||||  
 Db 505 CAGGCAATCCGATTCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 564  
 QY 121 ValGluGluTyGlyLysAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
 |||||  
 Db 565 GTGCAAGAACTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 624  
 QY 141 AspLysLeuLysSerLysArgThrAspLeuLysAspHisAlaLeuHisArgSerHisAsp 160  
 |||||  
 Db 625 GACAAACTTTGAGTAAAGGAGACAGATCTTATGTGACCATGCTCGACATCTCAGAT 684  
 QY 161 GluLeu 162  
 |||||  
 Db 685 GAGCTA 690

RESULT 5  
 AB015631 814 bp mRNA linear PRI 26-JUL-2001  
 LOCUS AB015631 Homo sapiens mRNA for type II membrane protein, complete cds,  
 DEFINITION  
 ACCESSION AB015631  
 VERSION AB015631.1 GI:4586839  
 KEYWORDS type II membrane protein.  
 SOURCE Homo sapiens gastric adenocarcinoma CDNA to mRNA,  
 clone:lib:pkal-meta-1 clone:HP10390.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

VDRKTIOMSEFRINPDGSSVYEVPAARSEAHITELLEVCDMAKEYQIDPSTR  
 KNYRVVRSNGESSELDQIRIDSDISGTLKFACESIVEVEDELEFFSREADNVK  
 DKLSKRTDLCDAHRSDEL\*  
 BASE COUNT 248 a 250 c 315 g 204 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.6e-77 Length: 1017  
 Score: 847.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-082-502-21 (1-162) x AF186115 (1-1017)  
 QY 1 Argatgserglnabpdluhsicysglvalacysarvalaleuvalaspgluenglutrp 20  
 |||||||  
 DB 403 CGAAGGAGCCAAAGATCTACAGCTGGAGCTGGAGGCTGGTGGATGAGATTGAGAGTGG 462  
 |||||||  
 QY 21 Gluilealaaatgvalaaspproluyslthrlleglmetgliserpheargllesanpro 40  
 |||||||  
 DB 463 GAAATGGCCCGCGTGACCCCAAGAGACCATTCAGATGGGATGCTCCGATCATCAATCCA 522  
 |||||||  
 QY 41 Aspelyserserlnservalalgluvalprotyralaargserglualahlsleuthglu 60  
 |||||||  
 DB 523 GATGGCAGCCAGTCAGTTGTGGAGTACCTTATCCCGCTCAGAGGCCACCTCAGACAG 582  
 |||||||  
 QY 61 Leuleuglgluvalcysasparmetlglutrgylgluignlileaspproserthr 80  
 |||||||  
 DB 583 TTGCTTGAGGAGGTGTGTGACCCGAAATGAGAGTACGGGGAACGATTGACCTTTCTACC 642  
 |||||||  
 QY 81 Hlsarglasasnyrvalarvalalseraragsnlgylusersergluleasplu 100  
 |||||||  
 DB 643 CACCGCAAGACTACGATACGCTGCTGAGCCGGAATGAGAACTCAGTACGACTGACTTA 702  
 |||||||  
 QY 101 Glucllyllearglileaspserserlleserllyhrlleulysphealacysgluserlle 120  
 |||||||  
 DB 703 CAGGGCATTCCCAATTTGATCAGATATCAGCGGACCCCTCAAGTTGCTGTGAGAGCAATT 762  
 |||||||  
 QY 121 Valgluglutyrgluaspglualeuilegluphpeserargluualaspsasvallys 140  
 |||||||  
 DB 763 GTGGAAGATTCAGAGATGATGATTCGATTCCTTCAGAGAGGCTGACACAGCTTTAAA 822  
 |||||||  
 QY 141 Asplysleucyserserllyarqthraspleucysasphlsalaleuhsarqserhlsasp 160  
 |||||||  
 DB 823 GACAAACTTTCAGTAAAGCGACAGATCTATGTGACCATGCCCTGCACAGATTCACAGAT 882  
 |||||||  
 QY 161 Gluleu 162  
 |||||||  
 DB 883 GAGCTA 888  
 |||||||  
 RESULT 3  
 AC012013 158198 bp DNA linear HTG 17-JUL-2001  
 LOCUS  
 DEFINITION Homo sapiens chromosome 12 clone RP11-764L14, WORKING DRAFT  
 SQUENCE: 5 unordered pieces.  
 AC012013  
 AC012013 GI:14717292  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
 SOURCE  
 Homo sapiens.  
 ORGANISM  
 Homo sapiens.  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 158198)  
 REFERENCE  
 AUTHORS  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbarta,J.,  
 Benton,U., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
 Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhaq,C.,  
 Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carion,T.F.,  
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
 Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
 Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
 Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,  
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,  
 Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,  
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,  
 Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S.,  
 Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,  
 Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,  
 Hollins,B., Homs,F., Howard,S., Huber,J., Huijck,S., Hume,D.,  
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,  
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korah,J.,  
 Kovar,C., Kratovic,Z., Kuresh,A., Landry,N., Leal,B., Lewis,L.C.,  
 Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,D., Liu,W.,  
 Louisedge,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
 Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A.,  
 Martinez,E., Massey,E., Mawhinney,E., Mcleod,M.P., Meador,M.,  
 Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,  
 Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,  
 Ogun,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,  
 Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,  
 Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shooshitari,N.,  
 Slason,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,  
 Stone,H., Sutton,A., Swalek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,  
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,  
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalona,D., Vinson,R.,  
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
 Wellington,S., Williams,G., Williamson,A., Wiczek,R., Wooden,S.,  
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstock,G., and Gdbbs.R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 158198)  
 Worley,K.C.  
 Direct Submission  
 Submitted (19-OCT-1999) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 13, 2001 this sequence version replaced gi:14547727.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: HMW2  
 Center clone name: RP11-764L14  
 ----- Summary Statistics  
 Sequencing vector: M13; L08821  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 161278 bases at least Q40  
 Consensus quality: 170806 bases at least Q30  
 Consensus quality: 174843 bases at least Q20  
 Estimated insert size: 168803; sum-of-coverage estimation  
 Estimated insert size: 163288; agarose-ef estimation  
 Quality coverage: 6x in Q20 bases; agarose-ef estimation  
 Quality coverage: 5.8x in Q20 bases; sum-of-coverage estimation  
 -----  
 NOTE: Estimated insert size may differ from sequence length  
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
 NOTE: This is a 'working draft' sequence; it currently  
 consists of 5 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.  
 1 108229: contig of 108229 bp in length  
 \* 108230 108329: gap of unknown length  
 \* 108330 150246: contig of 41917 bp in length



BC008261  
LOCUS BC008261 798 bp mRNA linear ROD 07-AUG-2002  
DEFINITION Mus musculus, transmembrane protein 4, clone MGC:6853  
IMAGE:2650612, mRNA, complete cds.  
ACCESSION BC008261  
VERSION BC008261.1 GI:14198400  
KEYWORDS MGC.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 798)  
REFERENCE  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (22-MAY-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
CONTACT: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu Xia Deng Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amgdbcm.tmc.edu](mailto:amgdbcm.tmc.edu)  
Guanerene, P.H., Garcia, A.M., Lu, X., Huliy, S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAX Plate: 5 Row: b Column: 23  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 9903606.  
FEATURES  
source  
1..798  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="129,C57BL/6J,FVB/N"  
/clone="MGC:6853 IMAGE:2650612"  
/tissue\_type="Mammary tumor. Brca1-/fl: MMTV-Cre model. 10  
months old, gross tissue."  
/clone\_lib="NCI CGAP\_Mam3"  
/lab\_host="DH10B"  
/note="Vector: PCMV-SPORT6"  
115..663  
/codon\_start=1  
/product="transmembrane protein 4"  
/protein\_id="AAH08261.1"  
/db\_xref="GI:14198401"  
/db\_xref="locusID:56530"  
/translation="MKGWGWLALLGLVLTGWTAMARRSODLHCACRALVDELEWEIAR  
VDPKTIQSGSRINPDGOSVYEVPAENSEHLLEVCDDRKKEGEQIDPSTHR  
KNYRVVNSGENSELQGIIRIDISDTLAFACSEIYEVEDELIFFSREADNVK  
DKLSKRDLCDHALHRSDEL"  
BASE COUNT 227 a 182 c 228 g 161 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.93e-77 Length: 798  
Score: 847.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0  
US-10-082-502-21 (1-162) x BC008261 (1-798)  
QY 1 ArgArgSerGlnAspIeuNHScysGlyAlaCysArgAlaLeuValAspGluLeuGluTTP 20

|||||  
Db 175 CGAAGAGCCAGATCTACACGTGAGCTTCAGGCTCTGTGATCAATTAGAGTGG 234  
QY 21 GUAIIAAlaArgValAspProLysLysThrIleGluMetGlySerPheArgIleAsnPro 40  
Db 235 GAATATGGCCCGCGTCGACCCCAAGAGACCATTCACATGGATCTTCCGAATCAATCCA 294  
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluValAsnLeuThrGlu 60  
Db 295 GATGGCAGCAGTCACTTGTGGAGTACTTATGCCCGTCAGAGGCCACCTCACAGAG 354  
QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGluIleAsnProSerThr 80  
Db 355 TTGCTTGAGAGAGTGTGTGATACCAATGAAGAGAGTACGGGAGACGATTCACCTTAC 414  
QY 81 HIsArgLysAsnTyrValAlaArgValValSerArgAsnGlyGluSerSerGluLeuAspLeu 100  
Db 415 CACGCCAGAACTACGTACGCGCTGTGAGCCGGAATGGAAATCCAGTAACTAGACTTA 474  
QY 101 GInGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
Db 475 CAGGCATCGAATTCATCATCATATACAGCGGACCTCAAGTTCCGTGTGAGACATT 534  
QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
Db 535 GTGGAAGAATACAGAGATGAGCTTATCGAATCTCTCCAGAGAGCTGCACAACTTAA 594  
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysArgHisAlaLeuHisArgSerHisAsp 160  
Db 595 GACAACTTTGCACTAGACGACAGATCTATGTGACCATGCTCCTGCACAGATCTCACAGAT 654  
QY 161 GluLeu 162  
Db 655 GAGCTA 660  
RESULT 2  
AF186115  
LOCUS AF186115 1017 bp mRNA linear ROD 13-JAN-2000  
DEFINITION Mus musculus putative secreted protein zsig9 (zsig9) mRNA, complete  
cds.  
ACCESSION AF186115  
VERSION AF186115.1 GI:6014635  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
REFERENCE  
AUTHORS Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1017)  
Sheppard, P., Jelinek, L., Whitmore, T., Blumberg, H., Lehner, J. and  
O'Hara, P.  
Mus musculus putative secreted protein  
Unpublished  
2 (bases 1 to 1017)  
Sheppard, P., Jelinek, L., Whitmore, T., Blumberg, H., Lehner, J. and  
O'Hara, P.  
Direct Submission  
Submitted (14-SEP-1999) Biomolecular Informatics, ZymoGenetics,  
Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA  
FEATURES  
source  
1..1017  
/organism="Mus musculus"  
/db\_xref="dbEST:AA049839"  
/db\_xref="taxon:10090"  
/clone="IMAGE:476061"  
1..1017  
/gene="zsig9"  
/gene="zsig9"  
343..891  
/gene="zsig9"  
/codon\_start=1  
/product="putative secreted protein zsig9"  
/protein\_id="AAF01433.1"  
/db\_xref="GI:6014636"  
/translation="MKGWGWLALLGLVLTGWTAMARRSODLHCACRALVDELEWEIAR

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:24:04 ; Search time 1535.7 Seconds  
(without alignments)  
3070.029 Million cell updates/sec

Title: US-10-082-502-21  
Perfect score: 847  
Sequence: 1 RRSQDLHCACGACALVDELEW.....LCGRKRDLCDAHRSDEL 162

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame+ -p2n -model -DEV-xip  
-O/cgn2.1/USPTO.spool/US10082502/rnatc.24032003.135059.6547/app.query.fasta.1.1308  
-DB-GenEmbl -OFMT-fastap -SUPFLX-rge -MINMATCH-0.1 -LOOPEXT-0 -LOOPEXT-0  
-UNITS-bits -START-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST-45  
-DOCALLGN-200 -THR\_SCORE-pct -THR\_MAX-100 -THR\_MIN-0 -ALIGN-15 -MODE-LOCAL  
-OUTFMT-pio -NOR-ext -HEAPSIZE-500 -MINLEN-200000000  
-USER-US10082502.qcgn.1.1.4758.qrunat.24032003.135059.6547 -NCP-6 -ICPU-3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES-0 -WAIT -LONGLOG -DEV\_TIMECUT-120  
-MARN\_TIMECUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6 -FGAPEXT-7  
-YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_ov.\*  
5: gb\_pat.\*  
6: gb\_ph.\*  
7: gb\_pl.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sts.\*  
11: gb\_sy.\*  
12: gb\_un.\*  
13: gb\_vl.\*  
14: gb\_vl.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_mu.\*  
21: em\_mu.\*  
22: em\_ov.\*  
23: em\_ov.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vl.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rnd.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	847	100.0	798	10	BC008261 Mus muscu
2	847	100.0	1017	10	AF186115 Mus muscu
3	847	100.0	158198	2	AC012013 Homo sapi
4	830	98.0	814	6	AX440456 Sequence
5	830	98.0	814	9	AB015631 Homo sapi
6	830	98.0	1080	9	AY032624 Homo sapi
7	830	98.0	1210	6	AX464040 Sequence
8	474.5	56.0	155023	2	AC025574 Homo sapi
9	472	55.7	182350	2	AC073896 Homo sapi
10	472	55.7	184762	2	AC023500 Homo sapi
11	451.5	53.3	207424	10	AC090489 Genomic s
12	429	50.6	101666	2	AC109891 Rattus no
13	253	29.9	649	9	AF186113 Homo sapi
14	253	29.9	824	9	BC001027 Homo sapi
15	248	29.3	2378	8	AK094445 Homo sapi
16	228.5	27.0	69753	2	AC103156 Rattus no
17	204	24.1	996	3	AY094698 Drosophila
18	193.5	22.8	40824	3	U13070 Caenorhabdi
19	156.5	18.5	622	5	AB055671 Danto rer
20	155	18.3	785	8	AY088187 Arabidops
21	154	18.2	564	6	AX440375 Sequence
22	141.5	16.7	175963	3	AC009379 Drosophila
23	141.5	16.7	279530	3	AE003518 Drosophila
24	140	16.5	55359	2	AC012854 Drosophila
25	140	16.5	172904	3	AC007414 Drosophila
26	140	16.5	275390	3	AE003831 Drosophila
27	138.5	16.4	1286	3	AY060987 Drosophila
28	133	15.7	161566	9	AC008060 Homo sapi
29	132	15.6	744	6	AX142811 Sequence
30	132	15.6	746	6	AX136616 Sequence
31	132	15.6	747	6	BD006701 Novel pol
32	132	15.6	1413	9	BC019903 Homo sapi
33	132	15.6	1419	6	AX136423 Sequence
34	132	15.6	1470	6	AX376412 Sequence
35	132	15.6	1470	6	AX454716 Sequence
36	132	15.6	1470	6	AX491194 Sequence
37	132	15.6	1512	6	AX142810 Sequence
38	132	15.6	1547	9	BC032339 Homo sapi
39	127	14.6	76187	2	AC015219 Drosophila
40	117	13.8	168418	9	AC009403 Homo sapi
41	116	13.7	1260	6	AX380553 Sequence
42	116	13.7	181896	2	AC114126 Rattus no
43	116	13.7	201649	2	AC120813 Rattus no
44	115.5	13.6	201649	2	AC120813 Rattus no
45	113	13.3	210631	2	AC129184 Mus muscu

RESULT 1

ALIGNMENTS



Db 165 GAAATGCCCCGGTGACCCCAAGACCAATTCAGATGGATCCTCCGAATCAATCCA 224  
QY 41 AspglySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
|||  
Db 225 GATGGCAGCCAGTGTGTGTGAGGTACTTATGCCCCGTGAGAGGCCACCTCCAGAG 284  
QY 61 LeuleuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
|||  
Db 285 TTGCTTGAGGAGGTGTGTGACCGAATGAAAGAGTACGGGAA-CAGATTGA-CCTTCTACC 342  
QY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 100  
|||  
Db 343 CACCGCAAGAACTACGTAACGCTGCTGAGCCGGAATGGAGATCCAGTGAACCTAGACTTA 402  
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
|||  
Db 403 CAGGCAATCCGAAATTGACTCAGATATCANGGCAACCCCTCAAGTTGCGTGTGAGAGCAT 462  
QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
|||  
Db 463 GTGGAAGAAATACGAGGATGACTTATCGAATCTCTCCAGAGAGGCTGACACACGTAAA 522  
QY 141 AsplysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSer-HisAs 160  
|||  
Db 523 GACAAACTTTGCAGTAACCGACAGATCTATGTGACCAATCCCTGCACAGATCCTCACGA 582  
QY 160 pGluLeu 162  
:|||||  
Db 583 AGAGCTC 589

Search completed: March 30, 2003, 22:43:01  
Job time : 130.151 secs



RESULT 13  
ABK36007  
ID ABK36007 standard; cDNA: 657 BP.  
XX  
XX  
AC ABK36007:  
XX  
XX  
DT 08-MAY-2002 (first entry)  
XX  
XX  
DE cDNA sequence #398 encoding novel human secreted protein.  
XX  
XX  
KM Human secreted protein: hyperproliferative disorder; autoimmune disorder;  
KM immune deficiency disorder; blood disorder; inflammatory disorder;  
KM infectious disorder; allergic condition; neurodegenerative disorder;  
KM liver fibrosis; coagulation disorder; gene therapy; antimicrobial;  
KM tumour; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN W0200177289-A2.  
XX  
PD 18-OCT-2001.  
XX  
PF 29-MAR-2001; 2001WO-US10232.  
XX  
PR 06-APR-2000; 2000US-195605P.  
XX  
PA (GEMV ) GENETICS INST INC.  
XX  
PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
PI Meberg D, Treacy M, Agostino MJ, Bowman MR, Spaulding V, Wong GC;  
PI Clark HF, Reichel K, Howes SH, Resnick RJ, Gulkota K, Graham JR;  
XX  
DR WPI: 2002-179322/23.  
XX  
XX  
PT Six hundred and twenty three polynucleotides derived from a variety of  
PT human tissue sources which encode secreted proteins, useful for  
PT treating immune deficiencies and disorders such as autoimmune disorders  
PT  
XX  
PS Claim 1: Page 297; 393pp; English.  
XX  
XX  
CC The present invention relates to the isolation of novel cDNA sequences  
CC which encode human secreted proteins. The cDNA sequences have been  
CC derived from a variety of human tissues. The invention also provides  
CC a method for producing proteins from these polynucleotide sequences.  
CC The proteins are useful for identifying compounds that modulate their  
CC activity and production. The sequences of the invention are  
CC useful for treating diseases such as hyperproliferative disorders  
CC (e.g. cancer), immune deficiency disorders (e.g. severe combined  
CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple  
CC sclerosis), blood disorders (e.g. thrombocytopenia), inflammatory  
CC disorders (e.g. arthritis), infectious disorders (e.g. hepatitis),  
CC allergic conditions (e.g. asthma), neurodegenerative disorders (e.g.  
CC Alzheimer's disease), liver fibrosis, coagulation disorders (e.g.  
CC hemophilia), and tumours. The polynucleotide sequences of the  
CC invention are also useful in gene therapy. ABK35610-ABK36232 represent  
CC the cDNA sequences of the invention that encode for novel human  
CC secreted proteins.  
XX  
SO Sequence 657 BP; 177 A; 149 C; 185 G; 146 T; 0 other;

Alignment Scores:  
Pred. No.: 2,49e-89 Length: 657  
Score: 815.00 Matches: 155  
Percent Similarity: 97.53% Conservative: 3  
Best Local Similarity: 95.68% Mismatches: 4  
Query Match: 96.22% Indels: 0  
DB: 24 Gaps: 0

US-10-082-502-21 (1-162) x ABK36007 (1-657)

OY 1 ArgArgSerGlnAspLeuHisCysGluAlaLeuArgAlaLeuValAspGluLeuGluTTP 20  
|||||

DB 49 CGAGAGACCCAGGATCTCCACTGTGGAGACATGACGGGCTTGTTGGATGAATCTGAATGG 108  
OY 21 GluIleAlaArgValAspProLysLeuThrIleGlnMetGlySerPheArgIleAspPro 40  
|||||  
DB 109 GAATTTGCCACGCGGAGACCCCAAGAGACATGACAGATGGATCTTCCGATCATTCGA 168  
OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
|||||  
DB 169 GATTGACACCCAGTACGTGTGTGGAGTGTCCCTATGCCCCCTCAGAGGCCCATCATCAGAG 228  
OY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGlyGlnIleAspProSerThr 80  
|||||  
DB 229 CTGCTGGAGGAGATATGTGACCGGATGAAGATATGGGGAACAGATTGATCTTCCACC 288  
OY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyGlySerGluLeuAspLeu 100  
|||||  
DB 289 CATGCAAGAACTACGTACGTGTAGTGGCCGGAATGAGATCAGAACTGGACCTA 348  
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
|||||  
DB 349 CAAGCATCCGAAATCGACTCAGATATTAGCCGACCCCTCAAGTTGCGTGTGAGAGCATTT 408  
OY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
|||||  
DB 409 GTGGAGGATATGAGGATGAACATGATTAATCTTTTCCAGAGGCTGACAAATGTTAA 468  
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
|||||  
DB 469 GACAAACTTTCAGATGAGCGAAGACAGATTTGTGTGACCATGCGCTGCACATATCCCATAT 528  
OY 161 GluLeu 162  
|||||  
DB 529 GAGCTA 534

RESULT 14  
AAK97837  
ID AAK97837 standard; cDNA: 832 BP.  
XX  
XX  
AC AAK97837:  
XX  
DT 23-SEP-1999 (first entry)  
XX  
XX  
DE Human secreted protein encoding cDNA #25.  
XX  
XX  
KM Secreted protein: human; cytosolic; thrombotic; osteopathic; forensic;  
KM diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
XX  
XX  
OS Homo sapiens.  
XX  
PN W09925825-A2.  
XX  
PD 27-MAY-1999.  
XX  
PF 13-NOV-1998; 98WO-IB01862.  
XX  
PR 04-SEP-1998; 98US-0099273.  
PR 13-NOV-1997; 97US-0066677.  
PR 17-DEC-1997; 97US-0069957.  
PR 09-FEB-1998; 98US-0074121.  
PR 13-APR-1998; 98US-0081563.  
PR 10-AUG-1998; 98US-0096116.  
XX  
XX  
PA (GEMT ) GENSET.  
XX  
PI Bougueleret L, Duclert A, Dumas Mline Edwards J;  
XX  
DR WPI: 1999-347472/29.  
XX  
DR P-PSDB; AAY36153.  
XX  
PT Extended cDNAs encoding secreted proteins  
XX  
XX  
XX Example 28; Page 194-195; 307pp; English.

CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
transgenic or knock out animals and can be used in gene therapy.

XX Sequence 1210 BP; 261 A; 316 C; 379 G; 254 T; 0 other;

## Alignment Scores:

Pred. No.:	8,51e-91	Length:	1210
Score:	830.00	Matches:	158
Percent Similarity:	98.77%	Conservative:	2
Best Local Similarity:	97.53%	Mismatches:	2
Query Match:	97.99%	Indels:	0
DB:	22	Gaps:	0

US-10-082-502-21 (1-162) x AAS21330 (1-1210)

```
OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
DB 603 CGAGAGAGCCAGATCTCCCTGTGGAGCATGCGAGGCTCTGTGGATGAACTAGAAATGG 662
OY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
DB 663 GAAATTGCCAGGTGACCCCAAGAACACCATTCAGATGGGATCTTCCGGATCAATCCA 722
OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
DB 723 GATGGCAGCCAGTCACTGGTGGGAGTGCCCTTATGCCGCTCAAGAGCCCACTCACAGAG 782
OY 61 LeuLeuGluGluValAlaCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
DB 783 CTCTGTGAGAGAGATATGTGACCGGATGAGAGATATGGGACAGATGATCTTCCACC 842
OY 81 HisArgLysAsnTyrValAlaArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 100
DB 843 CATCCCAAGAACTACTAGCTAGTGGGCGCGGAATGAGAAATCCAGTGAATCGACCTA 902
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 903 CAAGGCATCCGAATCGACTCAGATATTAGCGGACCCCTCAAGTTTCGTGTGAGACATT 962
OY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgLysAlaAspAsnValLys 140
DB 963 GTGAGAGATACAGAGATGAACTGATTTCTTTCCGAGAGGCTGCAATGTTAA 1022
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
DB 1023 GACAAACTTTGACGTAAAGCAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 1082
OY 161 GluLeu 162
DB 1083 GAGCTA 1088
RESULT 12
AAX97884
ID AAX97884 standard; cDNA; 714 BP.
AC AAX97884:
XX
XX 23-SEP-1999 (first entry)
XX
XX Human secreted protein encoding cDNA #72.
XX
XX Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;
XX diagnostic; gene therapy; chromosome mapping; secretion vector; ss.
XX
XX Homo sapiens.
XX
XX W09925825-A2.
XX
XX 27-MAY-1999.
XX
XX 13-NOV-1998; 98MO-IB01862.
XX
XX 04-SEP-1998; 98US-0099273.
```

PR 13-NOV-1997; 97US-0066677.  
PR 17-DEC-1997; 97US-0069957.  
PR 09-FEB-1998; 98US-0074121.  
PR 13-APR-1998; 98US-0081563.  
PR 10-AUG-1998; 98US-0096116.

XX (GEST ) GENSET.

XX Bouquellet L, Duclert A, Dumas Mline Edwards J;

XX WPI: 1999-347472/29.

XX F-PSDB; AAT36200.

XX Extended cDNAs encoding secreted proteins

XX Claim 1; Page 272; 307pp; English.

XX AAX97813-X97906 represent extended cDNA's which encode novel human

XX secreted proteins (see AAY36129-Y36222) and which have cytosolic,

XX thrombotic and osteopathic activity. The extended cDNAs can be used to

XX express secreted proteins or parts of them or to obtain antibodies

XX capable of binding to the secreted proteins. They may also be used in

XX diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX uses also include design of expression vectors and secretion vectors.

XX Sequence 714 BP; 203 A; 158 C; 202 G; 151 T; 0 other;

US-10-082-502-21 (1-162) x AAX97884 (1-714)

```
OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
DB 93 CGAGAGAGCCAGATCTCCCTGTGGAGCATGCGAGGCTCTGTGGATGAACTAGAAATGG 152
OY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
DB 153 GAAATTGCCAGGTGACCCCAAGAACACCATTCAGATGGGATCTTCCGGATCAATCCA 212
OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
DB 213 GATGGCAGCCAGTCACTGGTGGGAGTGCCCTTATGCCGCTCAAGAGCCCACTCACAGAG 272
OY 61 LeuLeuGluGluValAlaCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
DB 273 CTCTGTGAGAGAGATATGTGACCGGATGAGAGATATGGGACAGATGATCTTCCACC 332
OY 81 HisArgLysAsnTyrValAlaArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 100
DB 333 CATCCCAAGAACTACTAGCTAGTGGGCGCGGAATGAGAAATCCAGTGAATCGACCTA 392
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 393 CAAGGCATCCGAATCGACTCAGATATTAGCGGACCCCTCAAGTTTCGTGTGAGACATT 452
OY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgLysAlaAspAsnValLys 140
DB 453 GTGAGAGATACAGAGATGAACTGATTTCTTTCCGAGAGGCTGCAATGTTAA 512
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
DB 513 GACAAACTTTGACGTAAAGCAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 572
OY 161 GluLeu 162
DB 573 GAGCTA 578
```

CC polynucleotides, based on which tissues they are most highly expressed  
 CC in, and include developing products for the diagnosis or treatment of  
 CC cancer, tumours, developmental abnormalities and foetal deficiencies,  
 CC autoimmune diseases, lymphomas, Alzheimer's and cognitive disorders,  
 CC schizophrenia, immunological disorders, immune deficiency diseases  
 CC (AIDS), mood disorders, respiratory disorders, arthritis, asthma,  
 CC haematopoietic disorders, neural disorders, skeletal disorders,  
 CC osteoporosis, metabolic disorders, cardiovascular disorders, endocrine  
 CC disorders or gastrointestinal disorders. The polypeptides are also useful  
 CC for identifying their binding partners. The present sequence represents a  
 CC gene encoding a human secreted protein (see descriptor line for gene  
 CC number and clone identification).

SO Sequence 1180 BP; 258 A; 304 C; 363 G; 250 T; 5 other;

Alignment Scores:

Pred. No.: 8,24e-91 Length: 1180  
 Score: 830.00 Matches: 158  
 Percent Similarity: 98.77% Conservative: 2  
 Best Local Similarity: 97.53% Mismatches: 2  
 Query Match: 97.99% Indels: 0  
 Gaps: 0

US-10-082-502-21 (1-162) x AAX22112 (1-1180)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 Db 573 CGAGAGACCCAGATCTCCATCTGAGACATGACAGGCGCTGCTGATGACATGACATG 632  
 QY 21 GluLeuAlaArgValAspProlybSerThrLeuMetGlySerPheArgIleAsnPro 40  
 Db 633 GAAATTTGCCAGGTGGACCCCAAGAACCATTCAGATGGATCTTCCGATCATCA 692  
 QY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 Db 693 GATGGCACCAGTCACTGCTGGAGTGCCTTATGCCCCCTCAGAGGCCACTCTCAAG 752  
 QY 61 LeuLeuGluGluValCysAspArgMetIleGluTyrGlyGluGlnIleAspProSerThr 80  
 Db 753 CTGCTGGAGAGATATGATGACCGCATGAAGAGATATGGGAAACATGATTCCTCCACC 812  
 QY 81 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 100  
 Db 813 CATGCCAAGAACTACGTATGCTGTGTGGCCGGAATGGAGAAATCCAGTAACCTGACCTA 872  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 Db 873 CAAGGCATCCGAAATCGATCATATATTAGCGCACCCCTCAAGTTGCTGTGAGAGCAT 932  
 QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
 Db 933 GTGGAGGAATACGAGAGTAAGCACTCATTCATTTCCGAGAGCGTCAATGTAA 992  
 QY 141 AspLysLeuCysSerIleArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
 Db 993 GACAAACTTTCAGATGACGAAACAGATCTTGTGACCATGCTCCACATATATGCGATAT 1052  
 QY 161 GluLeu 162  
 Db 1053 GAGCTA 1058  
 RESULT 11  
 AAS21330  
 ID AAS21330 standard; cDNA; 1210 BP.  
 XX  
 AC AAS21330;  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human cDNA sequence encoding for PRO4426 polypeptide.  
 XX  
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
 KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;

KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
 KW adipocyte; A-peptide; factor VIIa; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO20010466-A2.  
 XX  
 PD 07-JUN-2001.  
 XX  
 PF 01-DEC-2000; 2000WO-US32678.  
 XX  
 PR 01-DEC-1999; 99WO-US28301.  
 PR 01-DEC-1999; 99WO-US28634.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 02-DEC-1999; 99WO-US28564.  
 PR 02-DEC-1999; 99WO-US28565.  
 PR 09-DEC-1999; 99US-0170262.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 20-DEC-1999; 99WO-US30911.  
 PR 20-DEC-1999; 99WO-US30999.  
 PR 30-DEC-1999; 99WO-US31243.  
 PR 06-JAN-2000; 2000WO-US00277.  
 PR 06-JAN-2000; 2000WO-US00376.  
 PR 11-FEB-2000; 2000WO-US03565.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 18-FEB-2000; 2000WO-US04342.  
 PR 22-FEB-2000; 2000WO-US04414.  
 PR 24-FEB-2000; 2000WO-US04914.  
 PR 24-FEB-2000; 2000WO-US05004.  
 PR 01-MAR-2000; 2000WO-US05601.  
 PR 20-MAR-2000; 2000WO-US07377.  
 PR 21-MAR-2000; 2000WO-US07532.  
 PR 30-MAR-2000; 2000WO-US08439.  
 PR 17-MAY-2000; 2000WO-US13705.  
 PR 22-MAY-2000; 2000WO-US14042.  
 PR 30-MAY-2000; 2000WO-US14941.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 10-NOV-2000; 2000WO-US30873.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,  
 PI Gerlitsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S,  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 XX  
 DR WPI: 2001-408281/43.  
 XX  
 PT P-PSDB; AAU12258.  
 XX  
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing  
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical -  
 XX  
 PS Claim 3; Fig 173; 813pp; English.  
 XX  
 CC AAS21244-AAS21518 encode for novel human secretory and transmembrane  
 CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, to modulate biological activities of cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from  
 CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC of T-lymphocytes, the release of a cytokine from peripheral blood  
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIa. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding





```
|||||
Db 265 GAAATGGCCAGGTGAGCCCAAGAACCATTCAGATGGGATCTTTCGGATCATCA 324
Oy 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
Db 325 GATGGCAGCAGTCAAGTGGTGGAGGTGCTTATGCCCGGTGAGAGCCCACTCAGAG 384
Oy 61 LeuLeuGluGluValCysAspArgMetCysGluTyrGlyGluGlnHisAspProSerThr 80
Db 385 CTGCTGGAGGAGATATGTGACCGGATGAAGAGTATGGGGAACACATTCCTCCACC 444
Oy 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGlnLeuAspLeu 100
Db 445 CATCCCAAGAACTACGTACGTAGTGGCCGGAATGGAGATCCAGTGAACCTGAGCTTA 504
Oy 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
Db 505 CAAGGCATCCGAATGACACTCAGATATATGCGGCACCCCAAGTTTGCCTGAGAGCAT 564
Oy 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
Db 565 GTGGAGGAAATACGAGGATGAACTCATTAATCTTTTCCCGAGAGCTGACAAATGTTAA 624
Oy 141 AspLysIleCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
Db 625 GACAAACTTTGACAGTAAAGCAAGATCTTTGTGACCATGCCCTGCACATATGCATGAT 684
Oy 161 GlnLeu 162
Db 685 GAGCTA 690

RESULT 8
ABL41995
ID ABL41995 standard; DNA; 814 BP.
XX
XX ABL41995;
XX
XX 11-JUN-2002 (first entry)
XX
XX Nucleotide sequence of human polypeptide HP10390.
XX
XX Antibody; antigen; transmembrane domain protein; HP10390; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH 145..693
XX FT /*tag= a
XX FT CDS /product= "HP10390"
XX
XX FN WO200208416-A1.
XX
XX PD 31-JAN-2002.
XX
XX PF 24-JUL-2001; 2001WO-JP06371.
XX
XX PR 24-JUL-2000; 2000JP-0222743.
XX PR 24-AUG-2000; 2000JP-0254407.
XX
XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX PI Kato S, Nagata N, Fujimura N, Kobayashi M, Ito K, Ishizuka Y;
XX
XX DR WPI; 2002-195877/25.
XX
XX DR P-PSDB; ABB09717.
XX
XX PT Antibody preparation by inoculation of an animal with a vector
XX expressing a fusion protein of an antigen on the C-terminal side of a
XX transmembrane domain for use as drugs, diagnostic reagents and
XX laboratory reagents
XX
XX PS Example; Page 33-35; 45pp; Japanese.
XX
```

```
CC The specification describes a method of antibody preparation. The
CC method comprises inoculating an animal with a vector expressing
CC a fusion protein having an antigen protein fused to the C-terminal
CC side (extracellular) of a transmembrane domain protein (the
CC N-terminal side of which is intracellular), and then isolating and
CC purifying the antibody from the animal. The antibodies can be used
CC as drugs, vaccines, diagnostic reagents and laboratory reagents. The
CC present sequence encodes a polypeptide, designated HP10390, which was
CC used in the course of the invention.
XX
SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;
Alignment Scores:
Pred. No.: 5.03e-91 Length: 814
Score: 830.00 Matches: 158
Percent Similarity: 98.77% Conservative: 2
Best Local Similarity: 97.53% Mismatches: 2
Query Match: 97.99% Indels: 0
DB: 24 Gaps: 0
US-10-082-502-21 (1-162) x ABL41995 (1-814)
Oy 1 ArgArgSerGlnAspLeuHisCysGlyValAcysArgAlaLeuValAspGluLeuTrp 20
Db 205 CGAGAGAGCCAGAGATCTCCACTGTCGACATGACAGGCTCTGGTGAATGAACTGAATGG 264
Oy 21 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40
Db 265 GAAATGGCCAGAGTGGAGCCCAAGAACCATTCAGATGGATCTTTCGGATCAATCA 324
Oy 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
Db 325 GATGGCAGCAGTCAAGTGGTGGAGGTGCTTATGCCCGGTGAGAGCCCACTCAGAG 384
Oy 61 LeuLeuGluGluValCysAspArgMetCysGluTyrGlyGluGlnHisAspProSerThr 80
Db 385 CTGCTGGAGGAGATATGTGACCGGATGAAGAGTATGGGGAACACATTCCTCCACC 444
Oy 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGlnLeuAspLeu 100
Db 445 CATCCCAAGAACTACGTACGTAGTGGCCGGAATGGAGATCCAGTGAACCTGAGCTTA 504
Oy 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
Db 505 CAAGGCATCCGAATGACACTCAGATATATGCGGCACCCCAAGTTTGCCTGAGAGCAT 564
Oy 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
Db 565 GTGGAGGAAATACGAGGATGAACTCATTAATCTTTTCCCGAGAGCTGACAAATGTTAA 624
Oy 141 AspLysIleCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
Db 625 GACAAACTTTGACAGTAAAGCAAGATCTTTGTGACCATGCCCTGCACATATGCATGAT 684
Oy 161 GlnLeu 162
Db 685 GAGCTA 690

RESULT 9
ABK09772
ID ABK09772 standard; cDNA; 814 BP.
XX
XX ABK09772;
XX
XX 14-MAR-2002 (first entry)
XX
XX DE Human ovarian tumour protein encoding cDNA #305.
XX
XX KW Human; ovarian tumour protein; cancer; cytostatic; immunostimulant; ss;
XX gene therapy; CD4+ T cell; CD8+ T cell; PCR primer.
XX
XX OS Homo sapiens.
XX
```

CC which encodes a 20 kd protein with one putative transmembrane  
 CC domain in the N-terminus. The cDNA was isolated from a human stomach  
 CC cancer cell line cDNA library. The protein has no homology with any  
 CC known proteins. The protein may be used to raise specific antibodies, as  
 CC assay reagents, as diagnostic tissue markers, for the isolation of  
 CC cognate receptors, ligands and binding proteins, and as biologically  
 CC active agents. Nucleotides encoding the protein may be used as primers  
 CC and probes or antisense molecules, and in gene therapy. Cells transformed  
 CC with these nucleotides may be used to screen for agonists and antagonists  
 CC which are potentially useful therapeutically.

XX  
 SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Alignment Scores:

Pred. No.:	5,03e-91	Length:	814
Score:	830.00	Matches:	158
Percent Similarity:	98.77%	Conservative:	2
Best Local Similarity:	97.53%	Mismatches:	0
Query Match:	97.99%	Indels:	0
DB:	21	Gaps:	0

US-10-082-502-21 (1-162) x AAK38327 (1-814)

QY	1	ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp	20
DB	205	CGAGAGAGCCAGATCTCCACTGTGAGAGCCGCTCTGTGATGACATGATGCG	264
QY	21	GluIleAlaArgValAspProLysThrIleGlnMetCysSerPheArgIleAsnPro	40
DB	265	GAATATGCCAGTGCAGCCCAAGAACACATTCAGATGGATCTTCCGATCAATCCA	324
QY	41	AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu	60
DB	325	GATGGAGCCAGCAGTGGGAGGAGTCCCTTATGCGCCGCTCAGAGCCCTCACAGAG	384
QY	61	LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr	80
DB	385	CTCTCTGAGAGCATATGTGACCGATGAGAGATGGGAGACAGATGATCCCTCCACC	444
QY	81	HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu	100
DB	445	CATCGCAAGAACTACGATCTGATGCGCCGGAATGGAAATCCAGTCAACATGACCTTA	504
QY	101	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle	120
DB	505	CAAGGATCCGATCGACATGATATTAGGCGACCTCAAGTTGGCTGTGAGAGCATT	564
QY	121	ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys	140
DB	565	GTCGAGAGATACAGAGATGAACTCATTTCTTTCCGAGAGGCTGACAAATGTTAAA	624
QY	141	AspLysLeuLysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp	160
DB	625	GACAAACTTTGCACTAAGCAGACAGATCTTTGACCATTCCTCCACATATGCGATGAT	684
QY	161	GluLeu 162	
DB	685	GACCTA 690	

RESULT 7

ID ABR52765 standard; cDNA; 814 BP.

XX ABR52765;

XX 15-AUG-2002 (first entry)

DE cDNA encoding transmembrane protein 4, a cancer-linked protein.

XX Expressed sequence tag; EST; human; cancer; anti-neoplastic;

KM cytotoxic T lymphocyte; chemotherapy; cytostatic; chromosome 12;

XX transmembrane protein 4; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 145..693

FT /tag "a

FT /product "Transmembrane protein 4"

PD 18-APR-2002.

PE 11-OCT-2001; 2001MO-US31607.

PR 11-OCT-2000; 2000US-239294P.

PR 11-OCT-2000; 2000US-239297P.

PR 11-OCT-2000; 2000US-239605P.

PR 12-OCT-2000; 2000US-239802P.

PR 12-OCT-2000; 2000US-239805P.

PR 12-OCT-2000; 2000US-240622P.

PR 19-OCT-2000; 2000US-241682P.

PR 19-OCT-2000; 2000US-241723P.

PR 31-OCT-2000; 2000US-244932P.

XX (AVAL-) AVALON PHARM.

XX Young PE, Horrigan S, Weaver Z, Endress GA;

XX WPI; 2002-463271/49.

XX P-PSDB; AA097063.

XX Identifying modulators of a cancer-related gene to screen agents for

XX preventing or treating cancer comprises detecting a difference in the

XX expression of cancer-linked genes in the presence or absence of test

XX compounds

XX Claim 1; Page 43; 66pp; English.

XX The invention relates to modulators of a cancer-related genes. Also

XX described are: (1) processes for identifying an anti-neoplastic agents

XX comprising contacting a cell exhibiting neoplastic activity with a

XX compound first identified as a cancer related gene modulator, and

XX detecting a decrease in the neoplastic activity. (2) a process for

XX determining the cancerous state of a cell by determining an increase in

XX the level of expression of at least one gene, where an elevated

XX expression relative to a known non-cancerous cell indicates a cancerous

XX state or potentially cancerous state. The anti-neoplastic agent is

XX useful for treating cancer or for protecting an animal against cancer.

XX The immunogenic composition is also useful for treating cancer in an

XX animal, where the composition elicits the production of cytotoxic T

XX lymphocytes specific for the immunogenic composition. Preferably, the

XX animal is a human. The cancer-linked genes and polypeptides are also

XX useful as targets for cancer therapy or chemotherapy. The present

XX sequence represents a cancer-linked gene located on chromosome 12,

XX which encodes transmembrane protein 4.

SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Alignment Scores:

Pred. No.:	5,03e-91	Length:	814
Score:	830.00	Matches:	158
Percent Similarity:	98.77%	Conservative:	2
Best Local Similarity:	97.53%	Mismatches:	0
Query Match:	97.99%	Indels:	0
DB:	24	Gaps:	0

US-10-082-502-21 (1-162) x ABR52765 (1-814)

QY	1	ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp	20
DB	205	CGAGAGAGCCAGATCTCCACTGTGAGAGCCGCTCTGTGATGACATGATGCG	264
QY	21	GluIleAlaArgValAspProLysThrIleGlnMetCysSerPheArgIleAsnPro	40

overexpression; antagonist; antibody; antisense nucleotide; tumour;  
 treatment; receptor; radio-label; fusion; polypeptide toxin; technique;  
 down-regulation; probe; diagnostic; therapeutic; cancer; brain; liver;  
 detection; stomach; lymphoma; alternative splicing; allelic variation;  
 silent mutation; ds.

XX Homo sapiens.  
 OS Synthetic.  
 XX

Key Location/Qualifiers  
 CDS 104..652  
 FT /tag- a  
 FT /product= "Zs19 secretory protein variant-4"  
 FT /note= "Overexpressed in tumours"  
 FT s19-peptide 104..163  
 FT /\*tag- b  
 FT mat-peptide 164..649  
 FT /\*tag- c  
 FT /label= "Mature-Zs19-protein-variant-4"

XX MO9960405-A1.  
 XX  
 XX  
 XX 25-NOV-1999.  
 XX  
 XX 19-MAY-1999; 99WO-US11107.  
 XX  
 XX 19-MAY-1998; 98US-0081183.  
 XX  
 XX (ZYMO ) ZYMOGENETICS INC.  
 XX  
 XX Moore EE, Taft DW;  
 XX WPI: 2000-039447/03.  
 XX DR P-PSDB; AAY15135.  
 XX

PT Detecting tumors using antibodies, antagonists and antisense  
 PT nucleotides to secretory protein-9 (Zs19) -  
 XX

PS Disclosure: Page 33-35; 45pp; English.

XX The present DNA sequence is a gene encoding the variant of the secretory  
 CC protein-9, Zs19 that arises due to alternative splicing, allelic  
 CC variation or silent mutations that result in amino acid changes. This  
 CC sequence is mapped to the human chromosome 12q15 region. It is  
 CC overexpressed in tumours. Antagonists, antibodies and antisense  
 CC nucleotides to Zs19 are useful for detecting and treating tumours. The  
 CC antagonist may be an antibody or receptor to Zs19 and it may be radio-  
 CC labelled or fused to a polypeptide toxin. It can be used for down  
 CC regulating the overexpression of Zs19. The gene sequence can be used as  
 CC nucleic acid probes to detect RNA encoding Zs19. The Zs19 sequence  
 CC facilitates improved diagnostic and therapeutic techniques for detecting  
 CC and treating cancers, especially of the brain, liver, stomach, lymphoma,  
 CC etc., at an early stage.  
 XX

SQ Sequence 806 BP; 229 A; 183 C; 233 G; 160 T; 1 other:

Alignment Scores:  
 Pred. No.: 4.97e-91 Length: 806  
 Score: 830.00 Matches: 158  
 Percent Similarity: 98.77% Conservative: 2  
 Best local Similarity: 97.53% Mismatches: 2  
 Query Match: 97.99% Indels: 0  
 DB: 21 Gaps: 0

US-10-082-502-21 (1-162) x AAZ08293 (1-806)

OY 1 ArgArgSerGlnAspLeuHisCysGluAlaCysArgAlaLeuValAspGluLeuGluThr 20  
 DB 164 CGGAGGAGCCGAGATCTCACTGTGAGCAGTCAAGGCTCTGTGGATGAATG 223  
 OY 21 GluLeuAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 40  
 DB 224 GAATTCGCCAGTGGACCCCAAGACACATTCAGATGGATCTTCCGGATCAATCA 283

OY 41 AspGlySerGlnSerValValGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 60  
 DB 284 GATGGCACCAGTCAGTGGTGGAGGTGCTTATGCCCTCAGAGGCCACCTCAGAG 343  
 OY 61 LeuLeuGluGluValCysAspArgMetLysGluGlyGluGlnIleAspProSerThr 80  
 DB 344 CTGCTGAGGAGATATGTGACCGGATGAGAGATGTGGGAAACAGATTGATCTTCCACC 403  
 OY 81 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 100  
 DB 404 CATGCCAAGACTACGTACGTCTACTGGCCCGAATGAGAAATCCAGTGAACCTGACCTTA 463  
 OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 DB 464 CAAGCATCTCGAATGACTCAGATATTAGCGGACCTCAAGTTGCGGTGAGAGCATT 523  
 OY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
 DB 524 GTGGAGGAATACGAGATGAACCTCATTAATCTTTTCCGAGAGGCTGACAAATGTTTAA 583  
 OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
 DB 584 GACAACTTTGCAGTACGAGAACAGATCTTGTGACCATGCCCTGCACATATCCGATGAT 643  
 OY 161 GluLeu 162  
 DB 644 GAGCTA 649

RESULT 6  
 AAZ38327  
 ID AAZ38327 standard; cDNA; 814 BP.  
 XX  
 XX AAZ38327;  
 XX  
 XX 09-FEB-2000 (first entry)  
 XX  
 XX Human transmembrane protein CDNA clone HP10390.  
 DE  
 XX HP10390; transmembrane domain; stomach cancer cell; antibody;  
 KW assay reagent; diagnostic marker; primer; probe; antisense; gene therapy;  
 KW agonist; antagonist; ligand; therapeutic; ds.  
 XX  
 XX Homo sapiens.  
 OS  
 XX

Key Location/Qualifiers  
 FT 145..693  
 FT CDS /tag- a  
 FT /product= "Human transmembrane protein HP10390"  
 FT

XX MO9955862-A2.  
 XX  
 XX 04-NOV-1999.  
 XX  
 XX 27-APR-1999; 99WO-JP02226.  
 XX  
 XX 28-APR-1998; 98JP-0119395.  
 XX  
 XX (SAGA ) SAGAMI CHEM RES CENT.  
 XX (PROT-) PROTEGENE INC.  
 XX  
 XX Kato S, Kimura T;  
 XX  
 XX WPI: 2000-023358/02.  
 XX DR P-PSDB; AAY52391.  
 XX  
 XX Human proteins with transmembrane domains, involved in control of cell  
 PT proliferation and differentiation, useful for treating e.g. cancer or  
 PT inflammation -  
 XX  
 XX Claim 4; Page 106-107; 114pp; English.  
 XX  
 CC This sequence represents the human CDNA clone HP10390

Db 301 CATCGAAGACTACGTACTGTATGGGCGGGAATGAGAAATCCAGTGAACCTGACCTA 360  
Qy 101 GINGLYILEARGYLEASPSERASPILESERGLYTHLEULYSPHEALACYSGLUSERILE 120  
Db 361 CAAGGATCCGAAATGACATCAGATATTAGCGGACCCCTCAAGTTGCGTGTGAGACATT 420  
Qy 121 VALGLUGLUTYRGILUASPGULUULEUGLUPHESERARGILUALAASPAENVALLYS 140  
Db 421 GTGAGAGAAATACGAGATCAACTGAAATCTTTCCCGAGAGGCTGCACATGTTAAA 480  
Qy 141 ASPLYSLEUCYSERLYSARGLTHASPLAUCYASPHISALALEUHSARGSERHISASP 160  
Db 481 GACAAACTTGGCAGTAAGCAACAGATCTTTGTGACCATGCTCGACATATGCGATGAT 540  
Qy 161 GLULEU 162  
Db 541 GAGCTA 546  
RESULT 4  
AA06969  
ID AAX06969 standard; cDNA; 806 BP.  
XX AAX06969;  
AC  
XX  
DT 10-MAY-1999 (first entry)  
XX  
DE Human secretory peptide-9 (Zs19) variant cDNA.  
XX  
XX Secretory peptide-9; Zs19; human; tumour marker; cancer; therapy;  
KM diagnosis; growth enhancer; variant; sg.  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 104..652  
FT FT /\*tag= a  
FT sig\_peptide 104..163  
FT FT /\*tag= b  
FT mat\_peptide 164..649  
FT FT /\*tag= c  
FT FT  
FT FT  
PN MO9901554-A1.  
PD 14-JAN-1999.  
XX  
PE 02-JUL-1998; 98MO-US13859.  
XX  
PR 17-JUN-1998; 98US-0099005.  
PR 03-JUL-1997; 97US-0051704.  
PR 03-JUL-1997; 97US-0888088.  
PR 19-MAY-1998; 98US-0081338.  
PR 19-MAY-1998; 98US-0085983.  
PR 17-JUN-1998; 98US-0089899.  
XX  
PA (ZYMO) ZYMOGENETICS INC.  
XX  
PI Jaspers SR, Jelinek LJ, Sheppard PO, Whitmore TE;  
XX  
DR MPI: 1999-106055/09.  
XX  
DR P-PSDB: AAW88474.  
XX  
PT New mammalian secretory peptide-9 (Zs19) - used as a growth  
XX PT enhancer for placenta, liver and heart, and as an indicator of  
XX PT cancer  
XX  
PS Claim 2: Page 73-74; 85pp; English.  
XX  
XX This cDNA clone encodes human secretory peptide-9, or Zs19,  
CC variant (see AAW88474). Zs19 (see also AAW88469) is overexpressed in  
CC human brain, liver, lung, oesophageal, stomach, colon, rectal,  
CC thyroid and lymphoma tumors. Thus, Zs19 can be used as an  
CC indicator for cancer. Zs19 cDNA was discovered in a placenta  
CC clone from a full-term pregnancy cDNA library which contained an

CC expressed sequence tag (see AAX06971). The invention provides  
CC polynucleotides (see AAX06968-70) encoding Zs19 polypeptides (see  
CC AAW88469-77) including mature polypeptides, other processed forms,  
CC variants and mouse orthologues. The Zs19 gene, or probes derived  
CC from it, can be used to determine if Zs19 is present on chromosome  
CC 12, and if a mutation has occurred. Antibodies raised against  
CC Zs19 can be used as diagnostic agents to determine the presence of  
CC Zs19, and thus the presence of cancer. They can also be labelled  
CC with radioisotopes or fused with toxins and used to treat tumours  
CC which overexpress Zs19. Antisense nucleotides derived from Zs19  
CC cDNA can also be used to inhibit the growth of tumour cells. Zs19  
CC proteins can be used to enhance the growth or development of the  
CC placenta, heart or liver.  
XX  
SQ Sequence 806 BP; 229 A; 183 C; 232 G; 160 T; 2 other;  
XX  
Alignment Scores:  
Pred. No.: 4.97e-91 Length: 806  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.99% Indels: 0  
DB: 20 Gaps: 0  
US-10-082-502-21 (1-162) x AAX06969 (1-806)  
Qy 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrp 20  
Db 164 CGAGAGACCGAGATCTCCCTGTGACATGCAGGGCTTGTGATGCACTGAAATGG 223  
Qy 21 GILILEALARGVALASPPROLYSLYSTRILEGLMETCYSERPHEARGILEANPRO 40  
Db 224 GAATATGCCAGGTGGACCCCAAGAACCATTCAGATGGATCTTCCGATCAATCCA 283  
Qy 41 ASPLYSERGLNSERVALVALIGLVALPROTYRALARGSERGLUALHISLEUTHGLU 60  
Db 284 GATGGAGCCAGTCAGTGGTGGAGTGCCTTATGCCGCTCAGAGGCCCTCCACAGAG 343  
Qy 61 LEULGUGLUGLVALCYSASPARGMETLYSGLUITYGLUGLILNILEASPPROSETRH 80  
Db 344 CTCTGGAGAGATATGTGACCGGATGAGAGATATGGGACAGATGATCTTCACC 403  
Qy 81 HISARGLYSANTRYVALARGVALVALSERARGASNGLYLSERSERGLUASPLEU 100  
Db 404 CATCGAAGACTACGTACTGTATGGGCGGGAATGAGAAATCCAGTGAACCTGACCTA 463  
Qy 101 GINGLYILEARGYLEASPSERASPILESERGLYTHLEULYSPHEALACYSGLUSERILE 120  
Db 464 CAAGGATCCGAAATGACATCAGATATTAGCGGACCCCTCAAGTTGCGTGTGAGACATT 523  
Qy 121 VALGLUGLUTYRGILUASPGULUULEUGLUPHESERARGILUALAASPAENVALLYS 140  
Db 524 GTGAGAGAAATACGAGATCAACTGATTCATTTCTTTCCGAGAGGCTGCACATGTTAAA 583  
Qy 141 ASPLYSLEUCYSERLYSARGLTHASPLAUCYASPHISALALEUHSARGSERHISASP 160  
Db 584 GACAAACTTGGCAGTAAGCAACAGATCTTTGTGACCATGCTCGACATATGCGATGAT 643  
Qy 161 GLULEU 162  
Db 644 GAGCTA 649  
RESULT 5  
AA08293  
ID AA08293 standard; DNA; 806 BP.  
XX  
XX AA08293;  
AC  
XX  
DT 07-FEB-2000 (first entry)  
XX  
DE Human Zs19 gene encoding secretory protein variant-4.  
XX  
KW Secretory protein-9; Human Zs19; chromosome 12q15 region; variant;

XX		The present DNA sequence is an ortholog encoding the secretory protein-9.
CC	Zs1g9	derived from mouse. It is overexpressed in tumours. Antagonists,
CC		antibodies and antisense nucleotides to Zs1g9 are useful for detecting
CC		and treating tumours. The antagonist may be an antibody or receptor to
CC	Zs1g9	and it may be radio-labelled or fused to a polypeptide toxin. It
CC		can be used for down regulating the overexpression of Zs1g9. The gene
CC		sequence can be used as nucleic acid probes to detect RNA encoding Zs1g9
CC		The Zs1g9 sequence facilitates improved diagnostic and therapeutic
CC		techniques for detecting and treating cancers, especially of the brain,
CC	Liver, stomach, lymphoma etc.,	at an early stage.
XX		
SQ	Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other:	
	Alignment Scores:	
	Pred. No.: 6,29e-93	Length: 1069
	Score: 847.00	Matches: 162
	Percent Similarity: 100.00%	Conservative: 0
	Best local Similarity: 100.00%	Mismatches: 0
	Query Match: 100.00%	Indels: 0
	DB: 21	Gaps: 0
US-10-082-502-21 (1-162) x AA208294 (1-1069)		
OY	1 ArgATSerGlnAspLeuHISCySGlyAlaCysArgAlaLeuValaspGluLeuGlnTrp	20
Db	418 CGAAGAGGCCCAAGTCTACACTGTGGAGCTTGAGGGCTCGTGATGAATAATTAGAGTGG	477
OY	21 GluIleuAlaArgValaAspProLysLyThrIleGluMeGlySerPheArgIleaAspro	40
Db	478 GAATTCGCCCGCGTGGAGACCACCAAGAAGACCATTCACAATGGANTCCYTCCGATTCATCCA	537
OY	41 AspGlySerGlnSerValValGluValProTyRAlaArgSerGluAlaHisLeuThrcLu	60
Db	538 GATGGCAGCCAGTCAAGTTGTGGAGGTACCTTAATGCCCGCTCAGAGGCCACCTCACAGAG	597
OY	61 LeuLeugluGluValCysAspArgMetLysGluTyrglGluGlnIleAspProSerThr	80
Db	598 TTGCTTGAGGAGGTGTGTGACCGAATGAAGAGTAGCGGGAACAGATTACCTTCTTACC	657
OY	81 HisArgLysAnTYrValaArgValaValserArgAsnGlyGluSerSerGluLeuAspLeu	100
Db	658 CACGCGAAGAACATACGATACCGTCGTGAGCGGAAATGGAGANATCCAGTGAACACTTA	717
OY	101 GlnGlyIleArgIleasPserAspIleSerGlyThrLeuLysPheAlaCysGluSerIle	120
Db	718 CAGGCGATCCGAAATTGACTCAGATATCAGGCGACCCCTCAAGTTGCCGTGAGAGCATT	777
OY	121 ValGluGluTyrgluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys	140
Db	778 GTGGAAACAATACGAGGTAGCTTATCGAATCTTCTCCAGAGAGGCTTACAAACCTTAA	837
OY	141 AspLysLeucySeriLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp	160
Db	838 GACAAACTTTGGCAGTAAAGCGAGACAGATCTATGTGACCAATGCCCTGCACAGATCTACGAT	897
OY	161 GluLeu 162	
Db	898 GAGCTA 903	
RESULT 3		
AAZ38326		
ID	AAZ38326 standard; cDNA; 546 BP.	
AC	AAZ38326;	
XX		
DT	09-FEB-2000 (first entry)	
XX		
DE	Human transmembrane protein cDNA clone HP10390 coding sequence.	
KW	HP10390; transmembrane domain; stomach cancer cell; antibody;	
KM	assay reagent; diagnostic marker; primer; probe; antisense; gene therapy;	
KW	agonist; antagonist; ligand; therapeutic; ds.	

XX	Homo sapiens.
OS	
XX	
XX	Key
XX	Location/Qualifiers
XX	1..540
FT	CDS
FT	/.*tag- a
FT	/product= "human transmembrane protein HP10390"
FT	/note= "No stop codon given in the specification"
XX	
XX	MO9955862-A2.
PN	
XX	
PD	04-NOV-1999.
XX	
PE	27-APR-1999; 99MO-JR02226.
XX	
PR	28-APR-1998; 98JP-0119395.
XX	
PA	(SAGA ) SAGAMI CHEM RES CENT.
PA	(PROT-) PROTEGENE INC.
XX	
P1	Kato S, Kimura T;
XX	
DR	WPI: 2000-023358/02.
DR	P-PSDB: AAY52391.
XX	
PT	Human proteins with transmembrane domains, involved in control of cell proliferation and differentiation, useful for treating e.g. cancer or inflammation -
PT	Inflammation
XX	
PS	Claim 3; Page 88; 11App; English.
XX	
CC	This sequence represents the coding sequence of human cDNA clone HP10390 which encodes a 20 kD protein with one putative transmembrane domain in the N-terminus. The cDNA was isolated from a human stomach cancer cell line cDNA library. The protein has no homology with any known protein. The protein may be used to raise specific antibodies, as assay reagents, as diagnostic tissue markers, for the isolation of cognate receptors, ligands and binding proteins, and, as biologically active agents. Nucleotides encoding the protein may be used as primers and probes or antisense molecules, and in gene therapy. Cells transformed with these nucleotides may be used to screen for agonists and antagonists which are potentially useful therapeutically.
CC	
CC	
CC	
CC	
CC	
SO	Sequence 546 BP; 139 A; 126 C; 165 G; 116 T; 0 other;
XX	
Alignment Scores:	
Pred. No.:	2,966-91 Length: 546
Score:	830.00 Matches: 158
Percent Similarity:	98.778 Conservative: 2
Best Local Similarity:	97.538 Mismatches: 2
Query Match:	97.998 Indels: 0
DB:	21 Gaps: 0
US-10-082-502-21 (1-162) x AA638326 (1-546)	
OY	1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
Db	CGAGAGGCCAGGATCTCCACCTGTGGAGCATGGAGGGCTCTGTGGAGAACTAACAATGG 120
OY	21 GluIleAlaArgValAspProLysIleThrIleGlnMetGlySerPheArgTlleasnPro 40
Db	GAATTTCCTCCAGGGGAGCCCCCAAGAACCAATTCAGATGGATCTTTCCGGATCAATCCA 180
OY	41 AspGlySerGlnSerValValIleGluValProTyraIlaArgSerGluAlaHISleuthrGlu 60
Db	GATGGCACGCCAGTCAGTGAGGGTGGCCCTTAATGCCCTCCAGAGGCCACCCTCCACAAG 240
OY	61 LeuLeuGluGluValCysAsparMetLysGluTyrgLyGluGlnIlaaspProserThr 80
Db	CTGTGGAGGAGATATGTGCACCGAATGAAGAGATATGGGAAACAATATGATCTTCCACC 300
OY	81 HisArgLysasnTyraArgValValSerArgasngLygIuseSerGluLeuAspLeu 100



1:	/SID52/gcgdata/genseq/genseqn-emb1/NA1980.DAT.*
2:	/SID52/gcgdata/genseq/genseqn-emb1/NA1981.DAT.*
3:	/SID52/gcgdata/genseq/genseqn-emb1/NA1982.DAT.*
4:	/SID52/gcgdata/genseq/genseqn-emb1/NA1983.DAT.*
5:	/SID52/gcgdata/genseq/genseqn-emb1/NA1984.DAT.*
6:	/SID52/gcgdata/genseq/genseqn-emb1/NA1985.DAT.*
7:	/SID52/gcgdata/genseq/genseqn-emb1/NA1986.DAT.*
8:	/SID52/gcgdata/genseq/genseqn-emb1/NA1987.DAT.*
9:	/SID52/gcgdata/genseq/genseqn-emb1/NA1988.DAT.*
10:	/SID52/gcgdata/genseq/genseqn-emb1/NA1989.DAT.*
11:	/SID52/gcgdata/genseq/genseqn-emb1/NA1990.DAT.*
12:	/SID52/gcgdata/genseq/genseqn-emb1/NA1991.DAT.*
13:	/SID52/gcgdata/genseq/genseqn-emb1/NA1992.DAT.*
14:	/SID52/gcgdata/genseq/genseqn-emb1/NA1993.DAT.*
15:	/SID52/gcgdata/genseq/genseqn-emb1/NA1994.DAT.*
16:	/SID52/gcgdata/genseq/genseqn-emb1/NA1995.DAT.*
17:	/SID52/gcgdata/genseq/genseqn-emb1/NA1996.DAT.*
18:	/SID52/gcgdata/genseq/genseqn-emb1/NA1997.DAT.*
19:	/SID52/gcgdata/genseq/genseqn-emb1/NA1998.DAT.*
20:	/SID52/gcgdata/genseq/genseqn-emb1/NA1999.DAT.*
21:	/SID52/gcgdata/genseq/genseqn-emb1/NA2000.DAT.*
22:	/SID52/gcgdata/genseq/genseqn-emb1/NA2001A.DAT.*
23:	/SID52/gcgdata/genseq/genseqn-emb1/NA2001B.DAT.*
24:	/SID52/gcgdata/genseq/genseqn-emb1/NA2002.DAT.*

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

[illegible]













Best Local Similarity: 18.78% Mismatches: 68  
 Query Match: 8.97% Indels: 56  
 DB: 1 Gaps: 7

US-10-082-502-21 (1-162) x US-07-872-644-38 (1-4131)

```

OY 6 LeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuGluTrpGluIleAlaArgVal 25
    ::::::::::::::::::::
DB 574 GTGCACCTGGT-----CAGCTGAGTGCATATGAGAGTGGAGCTGCACAGCTGTG 624
OY 26 AspProLys-----LysThrIleGluMetGlySerPheArg 37
    ::::::::::::::::::::
DB 625 GAGAAAGCATACCTGCTGGCCCTGAAAGAGGTGCAGGCCCTTCACAGCGGAGTCCAGC 684
OY 38 IleAsnProAspGlySerGlnSerValAlaGlu----- 48
    ::::::::::::::::::::
DB 685 GTGGCCCGGAGAGCAGCCAGATCCCGAGAGGAGCGAGCGGAGACCAAGAGGTGGG 744
OY 49 ValProTyrAlaArgSerGluAlaHisLeuThrGluLeuGluGluValCysAspArg 68
    ::::::::::::::::::::
DB 745 GTCCGATACACAAACCAAGACGAAAGATCTGACAGCTTGGCGGAGAGCTTCACGAC--- 801
OY 69 MetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValArgVal 88
    ::::::::::::::::::::
DB 802 -----CTGCATGCATCTTCCCTGCACCTCAAGTCTCCCAATAT 840
OY 89 ValSerArgAsnGlyLysSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 108
    ::::::::::::::::::::
DB 841 CTGCACCAAGAGAGACCCAGCATCCGCTGCTGCTGCTGCTGATCCGAGAGCAAT--- 897
OY 109 IleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGluTyrGluAspGluLeu 128
    ::::::::::::::::::::
DB 898 -----CTTCACCTCTCTGCAAGTGCATGAGTAAGTACTGAGAGAGAG 945
OY 129 IleGluPhe-----PheSerArgGluAlaAspAsnValLysAspLysLeu----- 143
    ::::::::::::::::::::
DB 946 ATCACCCTTCCGTTGACCAAGACGACCTGGGCCCAAGTGTGGAGACAAAGACTATC 1005
OY 144 -----CysSer 145
DB 1006 CAGCTGAAGAATCTCACCCTCGAGATATGCAAACAGCTGCAAAAGCATTTGGGCTGTAG 1065
OY 146 LysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAspGluLeu 162
DB 1066 GTGCAGGCCATGCTCTGTCTCTGTCATCAGCGGCGCAGTGCAGGCTC 1116

RESULT 10
US-08-297-494-38
: Sequence 38, Application US/08297494
: Patent No. 5580771
: GENERAL INFORMATION:
: APPLICANT: Beavo, Joseph A.
: APPLICANT: Bentley, Kelley
: APPLICANT: Charbonneau, Harry
: APPLICANT: Sonnenburg, William K.
: TITLE OF INVENTION: DNA Encoding Mammalian
: NUMBER OF SEQUENCES: 58
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
: STREET: Two First National Plaza, 20 South Clark
: STREET: Street
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60603
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
  
```

```

: APPLICATION NUMBER: US/08/297,494
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/688,356
: FILING DATE: 04-APR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 5580771and, Greta E.
: REGISTRATION NUMBER: 35,302
: REFERENCE/DOCKET NUMBER: 27866/30822
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 346-5750
: TELEFAX: (312) 984-9740
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 38:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4131 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 148..2910
: US-08-297-494-38

Alignment Scores:
Pred. No.: 12.3 Length: 4131
Score: 76.00 Matches: 37
Percent Similarity: 37.06% Conservative: 36
Best Local Similarity: 18.78% Mismatches: 68
Query Match: 8.97% Indels: 56
DB: 1 Gaps: 7

US-10-082-502-21 (1-162) x US-08-297-494-38 (1-4131)
OY 6 LeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuGluTrpGluIleAlaArgVal 25
    ::::::::::::::::::::
DB 574 GTGCACCTGGT-----CAGCTGAGTGCATATGAGAGTGGAGCTGCACAGCTGTG 624
OY 26 AspProLys-----LysThrIleGluMetGlySerPheArg 37
    ::::::::::::::::::::
DB 625 GAGAAAGCATACCTGCTGGCCCTGAAAGAGGTGCAGGCCCTTCACAGCGGAGTCCAGC 684
OY 38 IleAsnProAspGlySerGlnSerValAlaGlu----- 48
    ::::::::::::::::::::
DB 685 GTGGCCCGGAGAGCAGCCAGATCCCGAGAGGAGCGAGCGGAGACCAAGAGGTGGG 744
OY 49 ValProTyrAlaArgSerGluAlaHisLeuThrGluLeuGluGluValCysAspArg 68
    ::::::::::::::::::::
DB 745 GTCCGATACACAAACCAAGACGAAAGATCTGACAGCTTGGCGGAGAGCTTCACGAC--- 801
OY 69 MetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValArgVal 88
    ::::::::::::::::::::
DB 802 -----CTGCATGCATCTTCCCTGCACCTCAAGTCTCCCAATAT 840
OY 89 ValSerArgAsnGlyLysSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 108
    ::::::::::::::::::::
DB 841 CTGCACCAAGAGAGACCCAGCATCCGCTGCTGCTGCTGCTGATCCGAGAGCAAT--- 897
OY 109 IleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGluTyrGluAspGluLeu 128
    ::::::::::::::::::::
DB 898 -----CTTCACCTCTCTGCAAGTGCATGAGTAAGTACTGAGAGAGAG 945
OY 129 IleGluPhe-----PheSerArgGluAlaAspAsnValLysAspLysLeu----- 143
    ::::::::::::::::::::
DB 946 ATCACCCTTCCGTTGACCAAGACGACCTGGGCCCAAGTGTGGAGACAAAGACTATC 1005
OY 144 -----CysSer 145
DB 1006 CAGCTGAAGAATCTCACCCTCGAGATATGCAAACAGCTGCAAAAGCATTTGGGCTGTAG 1065
OY 146 LysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAspGluLeu 162
  
```

QY 145 SerLysArgThrAspLeu 150  
:::|||||::: |||  
Db 956 ACTAAAAAGCTGATTG 973

RESULT 8  
US-08-961-527-208  
Sequence 208, Application US/08961527  
Patent No. 6420135  
GENERAL INFORMATION:  
APPLICANT: Charles Kunach  
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 391  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,527  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: P8340P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 208:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3752 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

US-08-961-527-208

Alignment Scores:  
Pred. No.: 9.23 Length: 3752  
Score: 76.50 Matches: 31  
Percent Similarity: 39.73% Conservative: 27  
Best Local Similarity: 21.23% Mismatches: 65  
Query Match: 9.03% Indels: 23  
Gaps: 4

US-10-082-502-21 (1-162) x US-08-961-527-208 (1-3752)

QY 15 ValAspGluLeuGluTrpGluIleAla-----ArgValAspProLysLysThr 30  
:::||||| |||:::|||||  
Db 701 ATTTCACCTTGGACAGAGAGTGGTACTGCTCAACATCACTAATTAATTGAAAAA 760  
:::|||||::: |||  
QY 31 IleGlnMetCysSerPheArgIleAsnProAspGlySerGlnSerValValGluValPro 50  
:::|||||::: |||  
Db 761 CTCTCTGCTGCTGCG-----GATCCTGATGATGCACAGAGTTATAGAACTAAA 811  
:::|||||::: |||

QY 51 TyrAlaArgSerGluAlaHisLeuThrGluLeuGluGluValLysAspArgMetLys 70  
:::||||| |||:::|||||  
Db 812 TTAATAAAAGAGAGAGCTGAGCTAAACGCTAAACAGCTGAGTACCAAAAAAACAACA 871  
:::|||||::: |||

QY 71 GluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValArgValLysSer 90  
|||:::|||||  
Db 872 GAACCTGAAGAACTCTTGACACGCTTGATCCCTGAAGGTAAGCTCAGGATGAATTAGAT 931  
|||:::|||||

QY 91 ArgAsnGlyCysSerGluLeuAspLeuGlnGlyIleArgIleAspSerAspIleSer 110  
:::|||||::: |||  
Db 932 AAAGAGCAGAGAGAGAGCTAGTGGATTAATAAGCTGATGATCAATAAAGTGGCT 991  
:::|||||::: |||

QY 111 GlyThrLeuLysPheAlaCysGluSerIleValGlnGluTyrGluAspLeuIleGlu 130  
:::|||||::: |||  
Db 992 GATTTA-----GAAAAAGAAATTAAGTAACCTTGAA 1021  
:::|||||::: |||

QY 131 PhePheSerArgGluAlaAsp-----AsnValLysAspLysLeuCys 144  
|||:::|||||  
Db 1022 ATATTACTGGAGGGGCTGATCCTGAAGATGATGATGCTGCTGCTTCAATAATTAAGCT 1081  
:::|||||::: |||

QY 145 SerLysArgThrAspLeu 150  
:::|||||::: |||  
Db 1082 GCTAAAAAGCTGAGTTA 1099  
:::|||||::: |||

RESULT 9  
US-07-872-644-38  
Sequence 38, Application US/07872644  
Patent No. 5389527  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/872,644  
FILING DATE: 19920420  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5389527and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4131 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 148..2910

US-07-872-644-38

Alignment Scores:  
Pred. No.: 12.3 Length: 4131  
Score: 76.00 Matches: 37  
Percent Similarity: 37.06% Conservative: 36





DB 1492 AAAAAAATTAATCAC-----ACAAAAAACAAAAAATTAATTAATCTTGAA 1536  
OY 123 GAGTGTGlu-----AAGGluLeuIleGlu---rhepheserAAGlu 135  
DB 1537 GATTATGAAGAGTCAAAAAAGATTATGAAAGATTACTTGAAAAAATTTGAAATGAAA 1596  
OY 136 AAlaspaSn-----ValIlyAspLyLeuCySerLySarg 147  
DB 1597 TTTAATAAATAATTGACAAAGATGTCGATGATAAATAATTCAGTGCAGAA 1647  
RESULT 5  
US-08-588-985-1  
; Sequence 1, Application US/08588985  
; Patent No. 5777094  
; GENERAL INFORMATION:  
; APPLICANT: MICHIOYUKI MATSUDA et al.  
; TITLE OF INVENTION: CDNA OF DOCK180 GENE AND DOCK180 PROTEIN  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/588,985  
; FILING DATE: January 19, 1996  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ. ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6519 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHEICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE: spleen cell of homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 24..5619  
; US-08-588-985-1  
Alignment Scores:  
Pred. No.: 13.1 Length: 6519  
Score: 78.00 Matches: 44  
Percent Similarity: 40.11% Conservative: 29  
Best Local Similarity: 24.18% Mismatches: 66  
Query Match: 9.21% Indels: 44  
DB: 1 Gaps: 8  
US-10-082-502-21 (1-162) x US-08-588-985-1 (1-6519)  
OY 6 LeuHiscysGlyAlaCysArgAlaLeuValAspGluLeuGluIuTrpGluIleAlaArgVal 25

DB 4660 TTACAGACCGGCTCCTGCAGACACCCCTGAGCCCATGAAA-----AGATCG 4707  
OY 26 AspProLyLeuThrIleGlnMetGlySerPheArgIleAspProAspGlySerGlnSer 45  
DB 4708 AGAAGCTCAAGGACGTGATGCTTCCAGATGCTCTTTCTGGCCGAAGGATCAGAACTC 4767  
OY 46 ValValGluValProGlyAlaArgSerGluAlaIleLeuThrGluLeuGluGluVal 65  
DB 4768 ATG-----GAGACAAA-GTACGAGAGGACCTGAGGCCGCTC 4802  
OY 66 CysAspArgMet-----LysGluTyrgly 73  
DB 4803 CACGAGAGGATGAGGCTCTTCAAAAGCTGAGAAAGGTGAGAAAGAGTACGGC 4862  
OY 74 GluGlnIleAspProSer-----ThrIleArgLysAsnTyrgly 86  
DB 4863 GTCCGAAATCATGCCCTCAAGCTGATGATGATGAGAGGACGCCGCCCTCCATGAGTG 4922  
OY 87 ArgValValSerArgAsnGlyGlySerSerGluLeuAspLeuGlnGlyIle---ArgIle 105  
DB 4923 CGGTCTTCAGATGCTTCCATCCCGCCCTGTGTGTGGCTCTGTCTTCCCTC 4982  
OY 106 AspSerAspIleSer-----GlyThrLeuLysPheAlaCysGlySerIleVal 121  
DB 4983 TCATCGACACGACCCCTCCAGACGAGCTCCGAGGCTTCCCTGAGCTTCCCTG 5042  
OY 122 GluGluTyrgluAspGluLeuIleGluPhePheSerArgGluAlaAspArgVal---Lys 140  
DB 5043 CCAAG-----AAATGCACTCCAGGTCGACGACAAAGGAAAGAAAGAACACAAACATCAA 5084  
OY 141 AspLyLeuCySerLySargThrAspLeuCyAspHisAlaLeuHisArgSerHisAsp 160  
DB 5085 GATGACCTGGAGAAAGGAGAAAGAACACAAAGGAAAGAAAGAAAGGAAAGGAAAGGAAAG 5144  
OY 161 GluLeu 162  
DB 5145 GAGATA 5150  
RESULT 6  
US-08-971-988-1  
; Sequence 1, Application US/08971988  
; Patent No. 5786461  
; GENERAL INFORMATION:  
; APPLICANT: MICHIOYUKI MATSUDA et al.  
; TITLE OF INVENTION: CDNA OF DOCK180 GENE AND DOCK180 PROTEIN  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/971,988  
; FILING DATE: 17-NOV-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/588,985  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:

```
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; FILE REFERENCE: 11000.05001
; CURRENT FILING DATE: 2000-11-28
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Mouse
US-09-724-864-14

Alignment Scores:
Pred. No.: 0.00327 Length: 915
Score: 96.50 Matches: 40
Percent Similarity: 42.68% Conservative: 30
Best Local Similarity: 24.39% Mismatches: 65
Query Match: 11.39% Indels: 29
DB: 4 Gaps: 8

US-10-082-502-21 (1-162) x US-09-724-864-14 (1-915)
QY 6 LeuHscGysglYAlaCysArgAlaLeuValAspLeuLeuGluTTPGluIleAlaArgVal 25
DB 158 CTGGCGTGCATGCTGCTCGGGGCTGGCCCTCCAGATGGGGCAACGCTGGCCAAAGCA 217
QY 26 AspProLysThrIleGlnMetGlySerPheArgIleAsnProAspGlySerGlnSer 45
DB 218 GAGCGTAAATCTC-----ACTCCAGACCCAGT---GGA 250
QY 46 ValValAluValProTyrAlaArgSerGluAlaHisLeuThrGluLeuGluGluVal 65
DB 251 TTGCGAGGAGCGT-----AGTCAATCCACAGATGTCCTGACACGAC 298
QY 66 CysAspArg---MetLysGluTyrGlyGluGlnIleAsnProSerThrHisArgLysAsn 84
DB 299 TGCCTCAGAACTGCGGCGATCTATGA-----GTTCATGAAGTGAC 340
QY 85 TyrValArgValAlaSerArgAsnGlyLysSerGluLeuAspLeuGlnGlyIleArg 104
DB 341 CAGATGAAGGGTCTCAGAGGGCCAGACTTACGAAGGGCCAGAG-----CCAAGA 391
QY 105 IleAspSerAspIleSerGly-----ThrLeuLysPheAlaCysGluSer 119
DB 392 ATCAGCGTGAATGATTTCTGGGGGTCCCTGGCCCAATAGCGCTCTCCAAAGACGTTTCAC 451
QY 120 IleValGluGluTyr---GluAspGluLeuIleGluPheSerArgGluAlaAspAsn 138
DB 452 TACCTGGGTGATTTGGAGAGACCATCTATGACGCCCAAGGCCAAAGCGAAT 511
QY 139 ValLysAspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSer 158
DB 512 CTGAGAGCGGTCTCTGTGGGGGCAACCCATGAGCGCTGCTCAGAGATGTCGGCCACG 571
QY 159 HisAspLeuLeu 162
DB 572 AGAGAAAGAGCTT 583

RESULT 4
US-08-257-073-10
; Sequence 10, Application US/08257073
; Patent No. 576597
; GENERAL INFORMATION:
; APPLICANT: Paolletti, Enzo
; APPLICANT: de Talsne, Charles
; APPLICANT: Tine, John A.
; TITLE OF INVENTION: MALARIA RECOMBINANT POXYVIRUS VACCINE
```

```
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtiss, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue, 25th floor
; CITY: New York
; STATE: New York
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/257,073
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/075,783
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/852,305
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,183
; FILING DATE: 20-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2570
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5181 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-257-073-10

Alignment Scores:
Pred. No.: 0.654 Length: 5181
Score: 87.00 Matches: 36
Percent Similarity: 43.95% Conservative: 33
Best Local Similarity: 22.93% Mismatches: 58
Query Match: 10.27% Indels: 30
DB: 1 Gaps: 7

US-10-082-502-21 (1-162) x US-08-257-073-10 (1-5181)
QY 16 AspLeuLeuGluTTPGluIleAlaArgValAspProLysThrIleGlnMetGlySer 35
DB 1192 GATCACCCTGATTAAGATATTTATTAGAGAAAAATTAAGGATGATTAACCT 1251
QY 36 PheArgIleAsnProAspGlySerGlnSerValAlaGluValProTyrAlaArgSerGlu 55
DB 1252 AAATCACAAGATCTCTAGAAATCTGTTCATTAACCAAAAGTCTTATCCAAATGATTT 1311
QY 56 AlaHis-----LeuThrGluLeuGluGluValCys-----AspArgMetLys 70
DB 1312 GTATATCTCTTACCATCATGATATTCATTAATCTAGCTGCAGATAATGATAAAAT 1371
QY 71 GluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValArgValAlaSer 90
DB 1372 TCATATGCTGATTAATGAATCCTGATACCTAAGAAAAATTAAGAAAAATTAATTA 1431
QY 91 ArgAsnGlyGlu-----SerSerGluLeuAspLeuGlnGly 102
DB 1432 GATTAATAAGCAAAAAATATTCATTAATACATTAATAAAAAACAATGATTTGAAAGAA 1491
QY 103 IleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGlu 122
```

```
SEQUENCE CHARACTERISTICS:
LENGTH: 744 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..744
US-09-163-285-3

Alignment Scores:
Pred. No.: 6.6e-08 Length: 744
Score: 132.00 Matches: 41
Percent Similarity: 47.44% Conservative: 33
Best Local Similarity: 26.28% Mismatches: 54
Query Match: 15.58% Indels: 28
Gaps: 7

US-10-082-502-21 (1-162) x US-09-163-285-3 (1-744)
QY 8 CysGlyAlaCysArgAlaLeuValAspGluLeuGluTrpGluLeuAlaArgValAsp--- 26
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 112 TCGCAAGTGTCTAAGCTGCTGAGACAGACAGCTACAGCGGAACTGACCGCGTCA 171
QY 27 ProGlyLeuThrIleGlnMetGlySerPheArgIleAsnProAspGlySerGlnSerVal 46
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 172 TCTCGAGAGGTGCTGGAGCTGGGGCAGAGTGTCTG-----GATACAGCAGAGAGAG 222
QY 47 ValGluValProGlyAlaArgSerGluAlaHisLeuThrGluLeuGluGluValCys 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 AGACAGCTGCTTACAGCGCTTTCAGAGACAGAGCTGGAGAGCGCTTGAAGAAATTATGT 282
QY 67 AspArgMetLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 GACGGGATCTCGACTAT-----AGTGTACCGCTGAGCCGCAAGGCC 324
QY 85 TyrValArgValValSerArgAsnGlyGluSer-----SerGluLeuAspLeu 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 325 TCACTGAGATATGCCAAGGCTCAGAGTCAAGATGATGCAACTGAAAGGCTATGTCAG 384
QY 101 GlnGlyIleArgIleAspSerAspIleSer-----ArgGlyAsn 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 385 AAGGGGTGAAGGTGATCTGGGATCCCTCTGAGCTTTGGATGATGAGCCCAAGCTGGAG 444
QY 111 ---GlyThrLeuArgPheAlaCysGluSerIleValGluGluGluGluGluGluGlu 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 GTCACATACCTCAAGAACGAGCTGACACCATTTGGAGAGAGTTTGAAGCAATTGTGGA 504
QY 130 Glu---PhePheSerArgGluAlaAspAsnValIleAspIleLeuGluCys 144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 505 GACTGGTACTTCCACCATCAGAGAGAGCCCTACAAATAATTTCTCTGT 552

RESULT 2
US-09-163-285-1
; Sequence 1, Application US/09163285
; Patent No. 6204013
; GENERAL INFORMATION:
; APPLICANT: Rhodadoust, Mehran
; TITLE OF INVENTION: NOVEL MSP-5 PROTEIN AND NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/163,285
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/090,398
FILING DATE: June 24, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MNT-049
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1512 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 125..868
US-09-163-285-1

Alignment Scores:
Pred. No.: 1.82e-07 Length: 1512
Score: 132.00 Matches: 41
Percent Similarity: 47.44% Conservative: 33
Best Local Similarity: 26.28% Mismatches: 54
Query Match: 15.58% Indels: 28
Gaps: 7

US-10-082-502-21 (1-162) x US-09-163-285-1 (1-1512)
QY 8 CysGlyAlaCysArgAlaLeuValAspGluLeuGluTrpGluLeuAlaArgValAsp--- 26
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 TCGCAAGTGTCTAAGCTGCTGAGACAGACAGCTACAGCGGAACTGACCGCGTCA 295
QY 27 ProGlyLeuThrIleGlnMetGlySerPheArgIleAsnProAspGlySerGlnSerVal 46
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 296 TCTCGAGAGGTGCTGGAGCTGGGGCAGAGTGTCTG-----GATACAGCAGAGAGAG 346
QY 47 ValGluValProGlyAlaArgSerGluAlaHisLeuThrGluLeuGluGluValCys 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 347 AGACAGCTGCTTACAGCGCTTTCAGAGACAGAGCTGGAGAGCGCTTGAAGAAATTATGT 406
QY 67 AspArgMetLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 407 GACGGGATCTCGACTAT-----AGTGTACCGCTGAGCCGCAAGGCC 448
QY 85 TyrValArgValValSerArgAsnGlyGluSer-----SerGluLeuAspLeu 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 TCACTGAGATATGCCAAGGCTCAGAGTCAAGATGATGCAACTGAAAGGCTATGTCAG 508
QY 101 GlnGlyIleArgIleAspSerAspIleSer-----ArgGlyAsn 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 AAGGGGTGAAGGTGATCTGGGATCCCTCTGAGCTTTGGATGATGAGCCCAAGCTGGAG 568
QY 111 ---GlyThrLeuArgPheAlaCysGluSerIleValGluGluGluGluGluGluGlu 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 569 GTCACATACCTCAAGAACGAGCTGACACCATTTGGAGAGAGTTTGAAGCAATTGTGGA 628
QY 130 Glu---PhePheSerArgGluAlaAspAsnValIleAspIleLeuGluCys 144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 629 GACTGGTACTTCCACCATCAGAGAGAGCCCTACAAATAATTTCTCTGT 676

RESULT 3
US-09-724-864-14
; Sequence 14, Application US/09724864
; Patent No. 6380362
```

GenCore version 5.1.4\_P5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:33:58 ; Search time 35.3198 Seconds  
(without alignments)  
1406.624 Million cell updates/sec

Title: US-10-082-502-21  
Perfect score: 847  
Sequence: 1 RRSODHRCACALVDELEW.....LCSKRIDLCDHALRSHDEL 162

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=framed\_p2n.model -DEV=xlp  
-O=/cgn2\_1/USPRO.spool/US10082502/runat\_24032003\_135100\_6616/app-query.fasta.1.1308  
-DB=Issued\_Patents\_NA -OEM=fastap -SUFFIX=ini -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsun62 -TRANS=human40.cdd  
-LIST=45 -DOALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USBR=US10082502\_6C6G\_1.1.131\_@runat\_24032003\_135100\_6616 -NCPD=6 -ICPD=3  
-NO.XLPXY -NO.MMAP -LARGEQUERY -NEG\_SCORES=0 -NAT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA.\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/5C.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/5D.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/5E.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/5F.COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	132	15.6	744	4	US-09-163-285-3
2	132	15.6	1512	4	US-09-163-285-1
3	96.5	11.4	915	4	US-09-724-864-14
4	87	10.3	5181	1	US-08-257-073-10
5	78	9.2	6519	1	US-08-588-985-1
6	78	9.2	6519	1	US-08-971-988-1
7	77.5	9.1	1924	3	US-08-961-083-159
8	76.5	9.0	3752	4	US-08-961-527-208
9	76	9.0	4131	1	US-07-872-644-38
10	76	9.0	4131	1	US-08-297-494-38
11	76	9.0	4131	1	US-08-297-510-38
12	76	9.0	4131	1	US-08-479-532-38

13	76	9.0	4131	1	US-08-455-526-38	Sequence 38, Appl
14	76	9.0	4131	1	US-08-455-525-38	Sequence 38, Appl
15	76	9.0	4131	3	US-09-139-491-38	Sequence 38, Appl
16	76	9.0	4131	5	PCT-US92-03222-38	Sequence 38, Appl
17	76	9.0	5558	4	US-08-961-527-103	Sequence 103, App
18	75	8.9	5661	4	US-08-938-105-2	Sequence 2, Appl
19	74.5	8.8	1216	4	US-09-071-035-3	Sequence 3, Appl
20	74.5	8.8	1347	4	US-09-071-035-1	Sequence 1, Appl
21	74.5	8.8	4371	1	US-08-803-973-1	Sequence 1, Appl
22	74.5	8.8	4371	1	US-08-803-972-1	Sequence 9, Appl
23	74	8.7	918	3	US-08-937-271-9	Sequence 41, Appl
24	74	8.7	1640	4	US-08-676-444-1	Sequence 42, Appl
25	74	8.7	3789	1	US-07-872-644-42	Sequence 42, Appl
26	74	8.7	3789	1	US-08-297-494-42	Sequence 42, Appl
27	74	8.7	3789	1	US-08-297-510-42	Sequence 42, Appl
28	74	8.7	3789	1	US-08-479-532-42	Sequence 42, Appl
29	74	8.7	3789	1	US-08-455-526-42	Sequence 42, Appl
30	74	8.7	3789	1	US-08-455-525-42	Sequence 42, Appl
31	74	8.7	3789	3	US-09-139-491-42	Sequence 42, Appl
32	74	8.7	3789	5	PCT-US92-03222-42	Sequence 42, Appl
33	73.5	8.7	633	4	US-09-134-001C-578	Sequence 578, App
34	73.5	8.7	633	2	US-08-837-029-1	Sequence 1, Appl
35	73.5	8.7	4865	3	US-08-894-017-24	Sequence 24, Appl
36	73.5	8.7	5177	6	5352450-1	Patent No. 5352450
37	73.5	8.7	12127	4	US-08-961-527-148	Sequence 148, App
38	72.5	8.6	41708	4	US-09-470-512A-3	Sequence 3, Appl
39	72	8.5	640	3	US-08-961-083-33	Sequence 33, Appl
40	72	8.5	11309	4	US-08-961-527-108	Sequence 108, App
41	71.5	8.4	2088	4	US-09-458-481B-3	Sequence 3, Appl
42	71.5	8.4	4394	2	US-08-750-152A-1	Sequence 1, Appl
43	71.5	8.4	5084	1	US-08-306-681B-21	Sequence 21, Appl
44	71.5	8.4	5084	5	PCT-US93-06251-25	Sequence 25, Appl
45	71	8.4	601	4	US-09-556-877-42	Sequence 22, Appl

## ALIGNMENTS

RESULT 1  
US-09-163-285-3  
Sequence 3, Application US/09163285  
Patent No. 6204013  
GENERAL INFORMATION:  
APPLICANT: Rhodadoust, Mehran  
TITLE OF INVENTION: NOVEL MSP-5 PROTEIN AND NUCLEIC ACID MOLECULES  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/163,285  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/090,398  
FILING DATE: June 24, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MNI-049  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 3:

10

11

12

```
Db 603 CGAGGAGCCAGATCTCCACGTGAGACATGAGGCGCTGTGTGATGAATGAG 662
Qy 21 GUAUUAUAAGValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
Db 663 GAAATTGGCCAGCTGGACCCCAAGAACATTCAGATGGGATCTTCCGGATCAATCCA 722
Qy 41 AspGlySerGlnSerValValGluValProTyraIaArgSerGluAlaHisLeuThrGlu 60
Db 723 GATGGCACCCAGTCACTGTGGAGGTGCTTATGCCGCTCAGAGGCCACCCTCACAGAG 782
Qy 61 LeuLeuGluGluValCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 80
Db 783 CTGCTGGAGAGAGATATGTGACCGGATGAGAGAGATAGGGAGACAGATGATCTTCACCC 842
Qy 81 HisArgLysAsnTyrrValaArgValaIserArgAsnGlyGlnSerSerGluLeuAspLeu 100
Db 843 CATCCAGAGAACTAGCTAGCTAGCTAGTGGCCGCGAATGGAGATCCAGTGAATGAGACCTA 902
Qy 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlnSerIle 120
Db 903 CAAGGCATCCGAAATCGACTAGATATTTAGCGGCACCTCAAGTTTGCCTGTGAGAGCAT 962
Qy 121 ValGluGluTyrgluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
Db 963 GTGGAGGAATACGAGGATGAATCAATTAATCTTTCCCGAGAGGCTGACAAATGTAA 1022
Qy 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
Db 1023 GACAAACCTTTGCAGTACGAGAACACATCTTTGTGACCATGCCCTGCACATATCGCATGAT 1082
Qy 161 GluLeu 162
Db 1083 GAGCTA 1088
```

Search completed: March 31, 2003, 11:59:14  
Job time : 51.9767 secs

```
DB 843 CATCGCAGAACTACCTACTGTAGTGGCGCGGAATGGAAATCCAGTGAACCTGACCTA 902
      |||
QY 101 G|G|G|Y|L|E|A|G|T|L|E|A|S|P|S|E|A|S|P|I|E|S|E|G|Y|T|H|L|E|U|Y|S|P|H|E|A|L|A|C|Y|S|G|I|U|S|E|T|I|E 120
      |||
DB 903 CAAGGCTCCGAAATCGACTCGATGATATTAGGCGACCCCTCAAGTTGGTGTGAGACATT 962
      |||
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluLysAspAsnValLys 140
      |||
DB 963 GTGGAGGAATACAGAGATGAACCTGATGAACTCTTTCCCGAGAGGCTGACAACTGTTAA 1022
      |||
QY 141 AspLysLeuCySerLysArgThrAspLeuCyAspHisAlaLeuHisArgSerHisAsp 160
      |||
DB 1023 GACAACTTTCGAGTAAAGCAGACAGATCTTTGTGACCATGCCCCCTGACATATCGCATGAT 1082
      |||
QY 161 GluLeu 162
      |||
DB 1083 GAGCTA 1088

RESULT 14
US-10-142-419-173
; Sequence 173, Application US/10142419
; Publication No. US20030044945A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C244
; CURRENT FILING DATE: 2002-05-10
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Saplen
US-10-142-419-173

Alignment Scores:
Pred. No.: 3,41e-99 Length: 1210
Score: 830.00 Matches: 158
Percent Similarity: 98.77% Conservative: 2
Best Local Similarity: 97.53% Mismatches: 2
Query Match: 97.99% Indels: 0
DB: 9 Gaps: 0

US-10-082-502-21 (1-162) x US-10-142-419-173 (1-1210)
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyrP 20
      |||
DB 603 CGGAGGAGCCAGATCTCCACTGTGGAGCATGCGAGGCTCTGTGGATGAACTGAAATGG 662
      |||
QY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
      |||
DB 663 GAAATTCGCCAGGTGAGCAGCCCAAGAGACAACTTCAGATGGGATCTTCCGGAATCAATCA 722
      |||
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
```

```
DB 723 GATGGAGCCAGCTACGTGGTGGAGTGCCCTTATATGCCGCTCAGAGGCCACACAGAG 782
      |||
QY 61 LeuLeuGluGluValLysAspArgMetLysGluTyrGlyGluGlnIleLeuProSerThr 80
      |||
DB 783 CTCTCGAGAGATATGATGACCGGATGAAGAGATATGGGAAACAGATTAATCTTCCACC 842
      |||
QY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyGluSerSerGluLeuAspLeu 100
      |||
DB 843 CATCGCAGAACTACCTACTAGTGTAGTGGCGCGGAATGGAAATCCAGTGAACCTGACCTA 902
      |||
QY 101 G|G|G|Y|L|E|A|G|T|L|E|A|S|P|S|E|A|S|P|I|E|S|E|G|Y|T|H|L|E|U|Y|S|P|H|E|A|L|A|C|Y|S|G|I|U|S|E|T|I|E 120
      |||
DB 903 CAAGGCTCCGAAATCGACTCGATGATATTAGGCGACCCCTCAAGTTGGTGTGAGACATT 962
      |||
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluLysAspAsnValLys 140
      |||
DB 963 GTGGAGGAATACAGAGATGAACCTGATGAACTCTTTCCCGAGAGGCTGACAACTGTTAA 1022
      |||
QY 141 AspLysLeuCySerLysArgThrAspLeuCyAspHisAlaLeuHisArgSerHisAsp 160
      |||
DB 1023 GACAACTTTCGAGTAAAGCAGACAGATCTTTGTGACCATGCCCCCTGACATATCGCATGAT 1082
      |||
QY 161 GluLeu 162
      |||
DB 1083 GAGCTA 1088

RESULT 15
US-10-123-262-173
; Sequence 173, Application US/10123262
; Publication No. US20030049816A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C38
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Saplen
US-10-123-262-173

Alignment Scores:
Pred. No.: 3,41e-99 Length: 1210
Score: 830.00 Matches: 158
Percent Similarity: 98.77% Conservative: 2
Best Local Similarity: 97.53% Mismatches: 2
Query Match: 97.99% Indels: 0
DB: 9 Gaps: 0

US-10-082-502-21 (1-162) x US-10-123-262-173 (1-1210)
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyrP 20
```

|||||  
DB 1083 GAGCTA 1088

## RESULT 12

US-10-143-114-173

; Sequence 173, Application US/10143114  
; Publication No. US20030036180A1

## GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C211  
; CURRENT FILING DATE: 2002-05-09  
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-143-114-173

## Alignment Scores:

Pred. No.: 3,41e-99 Length: 1210  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.99% Indels: 0  
Gaps: 0  
DB: 9

US-10-082-502-21 (1-162) x US-10-143-114-173 (1-1210)

OY 1 Argatgserglnaspleuhsicysglvalacysargalaleuvalaspluleuglutr 20  
DB 603 CGAGGAGCCAGATCTCCAGTGTGAGACATGCAAGGCTGTGTGATGAATGAG 662  
OY 21 GluilealaargvalasprolyslsThrllleglmetglserphearglleaspro 40  
DB 663 GAAATTTGCCAGGTGGAGCCCAAGAAAGCAATTCAGATGGGATCTTTCGGATCAATCA 722  
OY 41 Aspglyserglnaservalvalgluvalprotyrllaargserglualahisleuthclu 60  
DB 723 GATGACACCGATCAGTGTGAGAGTGCCTTATGCCCTCAGAGCCACCTCACAGAG 782  
OY 61 Leuileugluvalucysaspargmetlysglytyrglygluglnlileasproserthr 80  
DB 783 CTGCTGGAGGAGATGTGACCGGATGAAGAGATGGGGAACAGATTGATCTTCCAC 842  
OY 81 HisarglysasnTYrvalargvalalseraraganglylusersergluleuasprou 100  
DB 843 CATCCCAAGAACTAGTACTGTACTGGCCGGAATGCAATCCAGTGAATGAGACTTA 902  
OY 101 GlncllylearglleaspsasrleaserglythrleuylsPhealacysglusertle 120  
DB 903 CAAAGGCAATCCGAAATGATGAGATGTAGCGGACCACTCAAGTTGGCTGTAGAGCAT 962  
OY 121 ValglugluTYrGluasprouleuilecluphepserargglualaspsasnvallys 140

|||||  
DB 963 GTGAGGAATACAGAGATGACTCATGTGATCTTTCCCGAGAGCTGACATGTTAAA 1022

## OY 141 AsplyleucysserlysrgrThraspleucysasprhlsalaleuhsilargserhlsap 160

DB 1023 GACAAACTTTGACATGACGAAACATATCTTTGTGACATGCCCTCAGATATGCGATGAT 1082

## OY 161 Gluleu 162

DB 1083 GAGCTA 1088

## RESULT 13

US-10-140-002-173

; Sequence 173, Application US/10140002  
; Publication No. US20030037623A1

## GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C59  
; CURRENT FILING DATE: 2002-05-06  
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-140-002-173

## Alignment Scores:

Pred. No.: 3,41e-99 Length: 1210  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.99% Indels: 0  
Gaps: 0  
DB: 9

US-10-082-502-21 (1-162) x US-10-140-002-173 (1-1210)

OY 1 Argatgserglnaspleuhsicysglvalacysargalaleuvalaspluleuglutr 20  
DB 603 CGAGGAGCCAGATCTCCAGTGTGAGACATGCAAGGCTGTGTGATGAATGAG 662  
OY 21 GluilealaargvalasprolyslsThrllleglmetglserphearglleaspro 40  
DB 663 GAAATTTGCCAGGTGGAGCCCAAGAAAGCAATTCAGATGGGATCTTTCGGATCAATCA 722  
OY 41 Aspglyserglnaservalvalgluvalprotyrllaargserglualahisleuthclu 60  
DB 723 GATGACACCGATCAGTGTGAGAGTGCCTTATGCCCTCAGAGCCACCTCACAGAG 782  
OY 61 Leuileugluvalucysaspargmetlysglytyrglygluglnlileasproserthr 80  
DB 783 CTGCTGGAGGAGATGTGACCGGATGAAGAGATGGGGAACAGATTGATCTTCCAC 842  
OY 81 HisarglysasnTYrvalargvalalseraraganglylusersergluleuasprou 100



```
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumes, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; PRIOR APPLICATION removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-474-173
```

```
Alignment Scores:
Pred. No.: 3,41e-99      Length: 1210
Score: 830.00           Matches: 158
Percent Similarity: 98.77%      Conservative: 2
Best Local Similarity: 97.53%    Mismatches: 2
Query Match: 97.99%           Indels: 0
DB: 9                     Gaps: 0
```

US-10-082-502-21 (1-162) x US-10-140-474-173 (1-1210)

```
OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrp 20
DB 603 CGAGAGAGCCAGATGTCCTGAGGATGCGAGGCTGTGGATGAGTGAACATGAG 662
OY 21 GlnIleAlaArgValAspProLysLysThrIleGlnMetCysPheArgIleAsnPro 40
DB 663 GAATATGCCAGGTGGACCCCAAGAGACCATTCAGATGGATCTTCCGATCAATCCA 722
OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
DB 723 GATGGCAGCCAGTCACTGAGTGGAGTCCCTTATGCCCGCTCAGAGGCCCTCAGAGAG 782
OY 61 LeuLeuGlnGluValCysAspArgMetLysGluTyrGlyGluGlnIleAsnProSerThr 80
DB 783 CTCTGGAGAGATATGTGACCGGATGAGAGATGCGGAGATGAGATTCCTCCACC 842
OY 81 HisArgLysAsnTyrValAlaArgValValSerArgAsnGlyLysSerGluLeuAspLeu 100
DB 843 CATCGCAAGACTACGTAAGTGTAGTGGCCGAGATGGAGATCCAGTGAACCTGAGACTA 902
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 903 CAAGGCTATCCGATCCGATCAGATATTAGCGGCACTTCAAGTTTGGCTGTGAGAGCAT 962
OY 121 ValGlnGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
DB 963 GTGAGAGAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATAA 1022
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
DB 1023 GACAACTTTGACAGTAAAGCAGACAGATCTTGTGACCATGCTCCCTGACATATGCGATGAT 1082
OY 161 GluLeu 162
DB 1083 GAGCTA 1088
```

RESULT 11  
US-10-142-431-173  
; Sequence 173, Application US/10142431  
; Publication No. US20030036179A1  
; GENERAL INFORMATION:

```
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumes, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-431-173
```

```
Alignment Scores:
Pred. No.: 3,41e-99      Length: 1210
Score: 830.00           Matches: 158
Percent Similarity: 98.77%      Conservative: 2
Best Local Similarity: 97.53%    Mismatches: 2
Query Match: 97.99%           Indels: 0
DB: 9                     Gaps: 0
```

US-10-082-502-21 (1-162) x US-10-142-431-173 (1-1210)

```
OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrp 20
DB 603 CGAGAGAGCCAGATGTCCTGAGGATGCGAGGCTGTGGATGAGTGAACATGAG 662
OY 21 GlnIleAlaArgValAspProLysLysThrIleGlnMetCysPheArgIleAsnPro 40
DB 663 GAATATGCCAGGTGGACCCCAAGAGACCATTCAGATGGATCTTCCGATCAATCCA 722
OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
DB 723 GATGGCAGCCAGTCACTGAGTGGAGTCCCTTATGCCCGCTCAGAGGCCCTCAGAGAG 782
OY 61 LeuLeuGlnGluValCysAspArgMetLysGluTyrGlyGluGlnIleAsnProSerThr 80
DB 783 CTCTGGAGAGATATGTGACCGGATGAGAGATGCGGAGATGAGATTCCTCCACC 842
OY 81 HisArgLysAsnTyrValAlaArgValValSerArgAsnGlyLysSerGluLeuAspLeu 100
DB 843 CATCGCAAGACTACGTAAGTGTAGTGGCCGAGATGGAGATCCAGTGAACCTGAGACTA 902
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 903 CAAGGCTATCCGATCCGATCAGATATTAGCGGCACTTCAAGTTTGGCTGTGAGAGCAT 962
OY 121 ValGlnGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
DB 963 GTGAGAGAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATAA 1022
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
DB 1023 GACAACTTTGACAGTAAAGCAGACAGATCTTGTGACCATGCTCCCTGACATATGCGATGAT 1082
OY 161 GluLeu 162
```

```

: SEQ ID NO 173
: LENGTH: 1210
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-176-921-173

```

## Alignment Scores:

```

Pred. No.: 3,41e-99      Length: 1210
Score: 830.00           Matches: 158
Percent Similarity: 98.77%      Conservative: 2
Best Local Similarity: 97.53%    Mismatches: 2
Query Match: 97.99%           Indels: 0
DB: 9                    Gaps: 0

```

US-10-082-502-21 (1-162) x US-10-176-921-173 (1-1210)

```

OY 1  ArgatSergInAspLeuHIScysGluAlaCysArgAlaLeuValAspGluLeuThrGlu 20
    |||||
DB 603 CGAGAGACCCAGAGATCTCCAGCTGTGAGACATGCGAGGCTCTGTGATGAACTAGAAATGG 662
OY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
    |||||
DB 663 GAAATGGCCAGAGTGCAGCCCAAGAGACCAATTCAGATGGAGATCTTCCGGATCAATCCA 722
OY 41 AspglySergInSerValValGluValProTyraAlaArgSergLualAlaHisLeuThrGlu 60
    |||||
DB 723 GATGGCACCCAGATGAGTGTGAGAGTGCCTTATGCCCCCTCAGAGGCCACCTCAGAGAG 782
OY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 80
    |||||
DB 783 CTGCTGGAGGAGATATGACCGGATGGAAGGATGGGGAACAGATGATCTTCCACC 842
OY 81 HisArgLysAsnTyraValArgValValSerArgAsnGlyGlySerSergLuleuAspLeu 100
    |||||
DB 843 CATCCCAAGAACTAGCTAGTGTACTGGCCGGAATGAGAAATCCAGATGAACTGGACCTA 902
OY 101 GlnGlyIleArgIleAspSergAspIleSergIlyThrLeuLysPheAlaCysGluSerIle 120
    |||||
DB 903 CAGGCAATCCGAAATGAGATCAATATTTAGCGCACCCCTCAAGTTGGCTGTGAGAGCAT 962
OY 121 ValGluGluTyrgluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
    |||||
DB 963 GTGGAGGAATAGAGAGATGAACTGATTAATCTTTCCGAGAGGCGACATGTTAAA 1022
OY 141 AsplysLeuCySserLysArgThrAspLeuCySAspHisAlaLeuHisArgSerHisasp 160
    |||||
DB 1023 GACAAACTTTCAGTAAGCAAGACAGATCTTTGTGACATGCCCTGCACATATCCGATGAT 1082
OY 161 GluLeu 162
    |||||
DB 1083 GAGCTA 1088

```

## RESULT 9

```

: Sequence 173, Application US/10137865
: Publication No. US20030032155A1

```

## GENERAL INFORMATION:

```

: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: Deforge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Fliviaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Collin K
: APPLICANT: Wood, William

```

```

: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P330R1C154
: CURRENT APPLICATION NUMBER: US/10/137,865
: CURRENT FILING DATE: 2002-05-03
: Prior Application removed - See Palm or File Wrapper
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 173
: LENGTH: 1210
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-137-865-173

```

## Alignment Scores:

```

Pred. No.: 3,41e-99      Length: 1210
Score: 830.00           Matches: 158
Percent Similarity: 98.77%      Conservative: 2
Best Local Similarity: 97.53%    Mismatches: 2
Query Match: 97.99%           Indels: 0
DB: 9                    Gaps: 0

```

US-10-082-502-21 (1-162) x US-10-137-865-173 (1-1210)

```

OY 1  ArgatSergInAspLeuHIScysGluAlaCysArgAlaLeuValAspGluLeuThrGlu 20
    |||||
DB 603 CGAGAGACCCAGAGATCTCCAGCTGTGAGACATGCGAGGCTCTGTGATGAACTAGAAATGG 662
OY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
    |||||
DB 663 GAAATGGCCAGAGTGCAGCCCAAGAGACCAATTCAGATGGAGATCTTCCGGATCAATCCA 722
OY 41 AspglySergInSerValValGluValProTyraAlaArgSergLualAlaHisLeuThrGlu 60
    |||||
DB 723 GATGGCACCCAGATGAGTGTGAGAGTGCCTTATGCCCCCTCAGAGGCCACCTCAGAGAG 782
OY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 80
    |||||
DB 783 CTGCTGGAGGAGATATGACCGGATGGAAGGATGGGGAACAGATGATCTTCCACC 842
OY 81 HisArgLysAsnTyraValArgValValSerArgAsnGlyGlySerSergLuleuAspLeu 100
    |||||
DB 843 CATCCCAAGAACTAGCTAGTGTACTGGCCGGAATGAGAAATCCAGATGAACTGGACCTA 902
OY 101 GlnGlyIleArgIleAspSergAspIleSergIlyThrLeuLysPheAlaCysGluSerIle 120
    |||||
DB 903 CAGGCAATCCGAAATGAGATCAATATTTAGCGCACCCCTCAAGTTGGCTGTGAGAGCAT 962
OY 121 ValGluGluTyrgluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
    |||||
DB 963 GTGGAGGAATAGAGAGATGAACTGATTAATCTTTCCGAGAGGCGACATGTTAAA 1022
OY 141 AsplysLeuCySserLysArgThrAspLeuCySAspHisAlaLeuHisArgSerHisasp 160
    |||||
DB 1023 GACAAACTTTCAGTAAGCAAGACAGATCTTTGTGACATGCCCTGCACATATCCGATGAT 1082
OY 161 GluLeu 162
    |||||
DB 1083 GAGCTA 1088

```

## RESULT 10

```

: Sequence 173, Application US/10140474
: Publication No. US20030032156A1

```

## GENERAL INFORMATION:

```

: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: Deforge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Fliviaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey

```

```
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
    |||
Db 603 CGAGAGAGCCAGCATCTCCACTGTGAGCATGACAGGGCTCTGTGATGAACTAGAAATGG 662
QY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetCysSerPheArgIleAsnPro 40
    |||
Db 663 GAAATTGCCAGGTGGACCCCAAGAAAGACCATTCAGATGGATCTTCCGGATCAATCCA 722
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
    |||
Db 723 GATGGAGCCAGCATGTGGTGGAGGTGCTTATAGCCCGCTCAAGGCCACACAGAG 782
QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
    |||
Db 783 CTCCTGAGAGAGATATCTGACCCGATGAAGAGATATGGGGAACAGATTGCTTCCACC 842
QY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 100
    |||
Db 843 CATCGCAAGAACTACGATGCTGTAGTGGCCGGAATGGAAATCCAGTGAATGGACCTA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
    |||
Db 903 CAAGGATCCGAATCGACTGAGATATTAGGGGACCCCTCAAGTTTGGCTGTAGAGCAT 962
QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
    |||
Db 963 GTGAGAGAAATACAGATGATCAATCTGATTTCCCGAGAGGCTGACAAATGTTAAA 1022
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
    |||
Db 1023 GACAAACTTTCAGATTAAGCAACAGATCTTGTGACCATCCCTGCACATATCGCATGAT 1082
QY 161 GluLeu 162
    |||
Db 1083 GAGCTA 1088

RESULT 7
US-10-176-918-173
; Sequence 173, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C382
; CURRENT APPLICATION NUMBER: US/10/176, 918
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-918-173

Alignment Scores: 3.41e-99 Length: 1210
Pred. No.:
```

```
Score: 830.00 Matches: 158
Percent Similarity: 98.77% Conservative: 2
Best Local Similarity: 97.53% Mismatches: 2
Query Match: 97.99% Indels: 0
DB: 9 Gaps: 0

US-10-082-502-21 (1-162) x US-10-176-918-173 (1-1210)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
    |||
Db 603 CGAGAGAGCCAGCATCTCCACTGTGAGCATGACAGGGCTCTGTGATGAACTAGAAATGG 662
QY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetCysSerPheArgIleAsnPro 40
    |||
Db 663 GAAATTGCCAGGTGGACCCCAAGAAAGACCATTCAGATGGATCTTCCGGATCAATCCA 722
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
    |||
Db 723 GATGGAGCCAGCATGAGGTGGAGGTGCTTATAGCCCGCTCAAGGCCACACAGAG 782
QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
    |||
Db 783 CTCCTGAGAGAGATATCTGACCCGATGAAGAGATATGGGGAACAGATTGCTTCCACC 842
QY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 100
    |||
Db 843 CATCGCAAGAACTACGATGCTGTAGTGGCCGGAATGGAAATCCAGTGAATGGACCTA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
    |||
Db 903 CAAGGATCCGAATCGACTGAGATATTAGGGGACCCCTCAAGTTTGGCTGTAGAGCAT 962
QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
    |||
Db 963 GTGAGAGAAATACAGATGATCAATCTGATTTCCCGAGAGGCTGACAAATGTTAAA 1022
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
    |||
Db 1023 GACAAACTTTCAGATTAAGCAACAGATCTTGTGACCATCCCTGCACATATCGCATGAT 1082
QY 161 GluLeu 162
    |||
Db 1083 GAGCTA 1088

RESULT 8
US-10-176-921-173
; Sequence 173, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C288
; CURRENT APPLICATION NUMBER: US/10/176, 921
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
```

Qy	81	HlslArGlysaantYrValAcgValValserArgAsnGlyGlubserGluLeuAspleu	100
Db	843	CATGCAGAAACTCGTACGTGTAGTGGCCGGGAATCGAATCCAGTGAACTCGACCTA	902
Qy	101	GlnClyYlIearGlyIeasPserAspIleSerGlyYthrLeuYsPheAlaCysGluSerIle	120
Db	903	CAAGGACATCCCAATCGCATGATATTAGCGGCAACCTCAAGTTTCGCTGTGAGACATT	962
Qy	121	ValGluGluYrGluAspGluLeuIleGluPhePheSerArgGluIuAlaAspAsnValYs	140
Db	963	GTCGAGCAATTCGAGGATGCAATCATTTGAAATTTCTTTTCCGAGAGGCTGACAAATGTTTAA	1022
Qy	141	AspIysLeuCysSerIysArgYthrAspLeuCysAspHisAlaLeuHisArgSerHisAsp	160
Db	1023	GACAAACTTTCACATTAAGCGCAACAGATCTTTGTGACCAATCCCTCGACATATTCGCATGAT	1082
Qy	161	GluLeu 162	
Db	1083	GAGCTA 1088	

OY 161 GluLeu 162  
| | | | |  
DB 1083 GACCTA 1088

RESULT 3  
US-10-121-049-173  
; Sequence 173, Application US/10121049  
; Publication No. US20030022239A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C17  
; CURRENT APPLICATION NUMBER: US/10/121,049  
; CURRENT FILING DATE: 2002-04-12  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-121-049-173

Alignment Scores:  
Pred. No.: 3,41e-99 Length: 1210  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.99% Indels: 0  
DB: Gaps: 0

US-10-082-502-21 (1-162) x US-10-121-049-173 (1-1210)

OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
| | | | |  
DB 603 CGGAGAGCCAGCATCTCCACTGTGGAGCAWCGCGCTGTGTGATGAATGAG 662

OY 21 GlnIleAlaArgValAspProLysThrIleGlnMetCysSerPheArgIleAspPro 40  
| | | | |  
DB 663 GAATATGCCAGATGGAGCCCAAGAACCATTCAGATGGAGATCTTCCGGAATCAATCA 722

OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
| | | | |  
DB 723 GATGGAGCCAGCATGAGTGGAGGTGCTTATGCGCGCTCAAGAGGCCCTCACAGAG 782

OY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
| | | | |  
DB 783 CTCCTGTGAGAGATATGTGACCGAGATGAGAGATGAGGAGACAGATTGCTTCCACC 842

OY 81 HisArgLysAspTyrValArgValValSerTyrAsnGlyLysSerGluLeuAspLeu 100  
| | | | |  
DB 843 CATCGAAGACTACTGCTAGTGTGAGGCCGGAATGAGAAATCCATGAACTGAGACTTA 902

OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
| | | | |  
DB 903 CAAAGCATCGAATCGACTGATGATATTAGGAGGAGACCCCTCAAGTTGGGTGTGAGAGCAT 962

OY 121 ValGluGluTyrGlnAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
| | | | |  
DB 963 GTGAGAGAAATACAGAGATGAATCATTCATTTCCCGAGAGGCTGACAAATGTTAAA 1022

OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
| | | | |  
DB 1023 GACAACTTTGACAGTAAAGCAAGATCTTGTGACCATCCCTGCACATATGCGATGAT 1082

OY 161 GluLeu 162  
| | | | |  
DB 1083 GACCTA 1088

RESULT 4  
US-10-123-904-173  
; Sequence 173, Application US/10123904  
; Publication No. US20030022328A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C54  
; CURRENT APPLICATION NUMBER: US/10/123,904  
; CURRENT FILING DATE: 2002-04-16  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-123-904-173

Alignment Scores:  
Pred. No.: 3,41e-99 Length: 1210  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.99% Indels: 0  
DB: Gaps: 0

US-10-082-502-21 (1-162) x US-10-123-904-173 (1-1210)

OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
| | | | |  
DB 603 CGGAGAGCCAGCATCTCCACTGTGGAGCAWCGCGCTGTGTGATGAATGAG 662

OY 21 GlnIleAlaArgValAspProLysThrIleGlnMetCysSerPheArgIleAspPro 40  
| | | | |  
DB 663 GAATATGCCAGATGGAGCCCAAGAACCATTCAGATGGAGATCTTCCGGAATCAATCA 722

OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
| | | | |  
DB 723 GATGGAGCCAGCATGAGTGGAGGTGCTTATGCGCGCTCAAGAGGCCCTCACAGAG 782

OY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
| | | | |  
DB 783 CTCCTGTGAGAGATATGTGACCGAGATGAGAGATGAGGAGACAGATTGCTTCCACC 842

PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069278  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069334  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069694  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 60/072320  
PRIOR FILING DATE: 1998-01-23  
PRIOR APPLICATION NUMBER: 60/073612  
PRIOR FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: 60/074086  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/074092  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-02-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081695  
PRIOR FILING DATE: 1998-04-14  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081818  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082999  
PRIOR FILING DATE: 1998-04-24  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085149  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/086414  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/086430  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088730  
PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088741  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090538  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07

Alignment Scores:  
Pred. No.: 3,41e-99  
Score: 830.00  
Percent Similarity: 98.77%  
Best Local Similarity: 97.53%  
Query Match: 97.99%  
DB: 9  
Gaps: 0

US-10-082-502-21 (1-162) x US-10-028-072-173 (1-1210)  
QY 1 ArgArgSerGlnAspPleuHlaGlyAlaGlyArgAlaLeuValAspGluLeuGluTrp 20  
DB 603 CGAGGAGGCCGATCTCCACCTGTGGAGCAGCGGCTGTGGATGAACCTGAATGG 662  
QY 21 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 40  
DB 663 GAAATTCGCCAGGTGGAGCCCAAGAACACCTTACATGATGGATCTTCCGATCAATCA 722  
QY 41 AspGlySerGlnSerValIleGluValProTyrIleAspSerGluAlaHlaLeuThrGlu 60  
DB 723 GATGGCAGCCAGTCAGTGTGGAGTGGCTTATGCCGCTCAGAGGCCACCTCAGAGAG 782  
QY 61 LeuLeuGluGluValCysAspArgMetIleGluTyrGluGlnIleAspProSerThr 80  
DB 783 CTGCTGGAGGAGATATGTGACCGGATGAAGATGGGAAAGATTGATGATCTTCCACC 842  
QY 81 HisArgLysAsnTyrValArgValIleSerArgAsnGlyLysSerGluLeuAspLeu 100  
DB 843 CATGCCAAGAACTACGATCGTACGTAGTGGCCGGAATGAGAACTCATGTAATGACCTA 902  
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 903 CAAGCATCCGAATCGACTCAGATATTAAGCGCACCTCAAGTTGGGTGTGACAGCATTT 962  
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
DB 963 GTGGAGGAATACGAGGATGAACCTCATTTCTTCCGAGAGGCTGACAAATGTAAA 1022  
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
DB 1023 GACAAATTTGCACTAAGCGAACAGATCTTGTGACCATGCCCTGCACATATTCGATGAT 1082



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 07:02:46 ; Search time 48.9767 Seconds  
(without alignments)  
2814.757 Million cell updates/sec

Title: US-10-082-502-21  
Perfect score: 847  
Sequence: 1 RNSQDHCACACALVDELM.....LCCKRIDCDHALRSHDEL 162

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 574371 segs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framed\_p2n.model -DEV=xlp  
-O=/cg2\_1/USPTO.spool/US10082502/runal\_24032003\_135103\_6870/app\_query.fasta.1.1308  
-DB=Published.Applications.NA -OPMT=fastap -SUFFIX=rmpb -MINMATCH=0.1  
-LOOPCU=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62  
-TRANS=human40.cdt -LIST=45 -DCCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAULEN=2000000000 -USER=US10082502.ecgn.1.1.145 -runat\_24032003\_135103\_6870  
-DEVT=6 -ICPU=3 -NO\_XLIPY -NO\_MMAL -LARGQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMECUT=120 -WARN\_TIMECUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published.Applications.NA.\*

1: /cg2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cg2\_6/ptodata/2/pubpna/PTCT\_NEW\_PUB.seq:\*  
3: /cg2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cg2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
5: /cg2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cg2\_6/ptodata/2/pubpna/PTCTUS\_PUBCOMB.seq:\*  
7: /cg2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cg2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cg2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cg2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cg2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cg2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cg2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cg2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	830	98.0	814	US-09-864-864-309	Sequence 309, App
2	830	98.0	1210	US-10-028-072-173	Sequence 173, App
3	830	98.0	1210	US-10-121-049-173	Sequence 173, App
4	830	98.0	1210	US-10-123-904-173	Sequence 173, App

5	830	98.0	1210	9	US-10-140-470-173	Sequence 173, App
6	830	98.0	1210	9	US-10-175-746-173	Sequence 173, App
7	830	98.0	1210	9	US-10-176-918-173	Sequence 173, App
8	830	98.0	1210	9	US-10-176-921-173	Sequence 173, App
9	830	98.0	1210	9	US-10-137-865-173	Sequence 173, App
10	830	98.0	1210	9	US-10-140-474-173	Sequence 173, App
11	830	98.0	1210	9	US-10-142-431-173	Sequence 173, App
12	830	98.0	1210	9	US-10-143-114-173	Sequence 173, App
13	830	98.0	1210	9	US-10-140-002-173	Sequence 173, App
14	830	98.0	1210	9	US-10-142-419-173	Sequence 173, App
15	830	98.0	1210	9	US-10-123-262-173	Sequence 173, App
16	830	98.0	1210	9	US-10-142-423-173	Sequence 173, App
17	830	98.0	1210	9	US-10-121-050-173	Sequence 173, App
18	830	98.0	1210	9	US-10-141-755-173	Sequence 173, App
19	815	96.2	657	9	US-09-822-846-198	Sequence 398, App
20	154	18.2	564	10	US-09-864-864-228	Sequence 228, App
21	132	15.6	1457	10	US-09-846-5738-110	Sequence 10, App
22	132	15.6	1470	9	US-10-035-885-58	Sequence 58, App
23	132	15.6	1470	9	US-10-035-885-58	Sequence 58, App
24	132	15.6	1470	9	US-10-174-590-479	Sequence 479, App
25	132	15.6	1470	9	US-10-176-758-479	Sequence 479, App
26	132	15.6	1470	9	US-10-175-737-479	Sequence 479, App
27	132	15.6	1470	9	US-10-173-706-479	Sequence 479, App
28	132	15.6	1470	9	US-10-175-738-479	Sequence 479, App
29	132	15.6	1470	9	US-10-175-752-479	Sequence 479, App
30	132	15.6	1470	9	US-10-176-482-479	Sequence 479, App
31	132	15.6	1470	9	US-10-176-757-479	Sequence 479, App
32	132	15.6	1470	9	US-10-176-913-479	Sequence 479, App
33	132	15.6	1470	9	US-10-180-552-479	Sequence 479, App
34	132	15.6	1470	9	US-10-180-557-479	Sequence 479, App
35	132	15.6	1470	9	US-09-931-836-58	Sequence 58, App
36	132	15.6	1470	9	US-10-173-700-479	Sequence 479, App
37	132	15.6	1470	9	US-10-174-572-479	Sequence 479, App
38	132	15.6	1470	9	US-10-174-579-479	Sequence 479, App
39	132	15.6	1470	9	US-10-174-582-479	Sequence 479, App
40	132	15.6	1470	9	US-10-174-588-479	Sequence 479, App
41	132	15.6	1470	9	US-10-175-739-479	Sequence 479, App
42	132	15.6	1470	9	US-10-175-740-479	Sequence 479, App
43	132	15.6	1470	9	US-10-175-743-479	Sequence 479, App
44	132	15.6	1470	9	US-10-176-488-479	Sequence 479, App
45	132	15.6	1470	9	US-10-176-492-479	Sequence 479, App

## ALIGNMENTS

RESULT 1  
US-09-864-864-309  
Sequence 309, Application US/09864864  
Patent No. US20020102679A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiansong  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Dillon, David C.  
APPLICANT: Secrist, Heather  
APPLICANT: Lodes, Michael J.  
APPLICANT: Algate, Paul A.  
APPLICANT: Fling, Steve P.  
APPLICANT: Mannion, Jane  
APPLICANT: Benson, Darin R.  
APPLICANT: Carter, Derrick  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.523  
CURRENT APPLICATION NUMBER: US/09/864, 864  
NUMBER OF SEQ ID NOS: 341  
SOFTWARE: Corixa Invention Disclosure Database  
SEQ ID NO 309  
LENGTH: 814  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-864-864-309



Query Match: 97.998 Indels: 0  
DB: 14 Gaps: 0

US-10-082-502-21 (1-162) x BM832828 (1-588)

```

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
    |||||||
DB 62 CGAGAGGAGCCAGATCTCCAGCTGTGGAGCATGAGGGCTCTGTGTGATGAACTAGAAATGG 121
    |||||||
QY 21 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40
    |||||||
DB 122 GAAATGGCCAGGTGAGCCCAAGAAAGACCATTACATGGGATCTTCCGATCAATCCA 181
    |||||||
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
    |||||||
DB 182 GATGGCAGCCAGTCACTGTGGAGTGCTTATGCCCGCTCAAGAGGCCCACTCACAGAG 241
    |||||||
QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
    |||||||
DB 242 CTGCTGGAGAGATATGTGACCGGATGAGAGATATGGGAGACAGATGTGATCCTTCACC 301
    |||||||
QY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 100
    |||||||
DB 302 CATCGCAAGACTACCTACGTGTAGTGGCCGGAATGAGAAATCCAGTGAATGAGACTG 361
    |||||||
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
    |||||||
DB 362 CAAGCATCCGATCGACTGATATTAGCGGACCCCTCAAGTTTCGTGTGAGAGCAT 421
    |||||||
QY 121 ValGluGluTyrGluAspGluLeuIleGlnPheSerArgGluLysAspAsnValLys 140
    |||||||
DB 422 GTGGAGGAATACAGAGATGAATCATTTCTTTCCCGAGAGGCTGACAAATGTTAAA 481
    |||||||
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
    |||||||
DB 482 GACAAACTTTCAGTAAGGACAGATGATCTTTGTGACCATGCCCTGCACATATCGCAT 541
    |||||||
QY 161 GluLeu 162
    |||||||
DB 542 GAGCTA 547
    |||||||

```

Search completed: March 31, 2003, 04:51:14  
Job time : 1041.65 secs

**JOURNAL  
COMMENT**

Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgrabs-remail.nih.gov  
This clone is available royalty-free through LNL; contact the  
IMAE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
GGI:954329

BASE COUNT	154 a	133 c	150 g	112 t
ORIGIN				

Alignment Scores:	
pred. No.:	6.28e-98
Score:	832.00
Percent Similarity:	98.77%
Best Local Similarity:	98.77%
Query Match:	98.23%
DB:	10
Length:	549
Matches:	160
Conservative:	2
Mismatches:	0
Indels:	0
Gaps:	0

US-10-082-502-21 (1-162) x BE198865 (1-549)

OY	1	ArgArgSerGluAspLeuHisCysGluValACysArgValAlaLeuValAspGluLeuGluTrp	20
Db	19	CGAAGGACCCAAGATCTTCACTGTGGAGCTTGCAGGGCTCTGGTGGATGAAATTAAGATGG	78
OY	21	GluIleAlaArgValAspProGlyLysThrIleGlnMetGlySerPheArgIleAsnPro	40
Db	79	GAAATTCGCCCGTGGACCCCAAGAACACCTTCAGATGGGATCCCTCCGATCAATCA	138
OY	41	AspIlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu	60
Db	139	GATGGCAGCCGCTAGTGTGTGGAGTACCTTATCCCGCTCAGAAAGCCACCTCACAGAG	198
OY	61	LeuLeuGluGluValLCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr	80
Db	199	TTCGTTAGAGAGGCTGTGTGACCCAGTAAGAGAGTACGGGGACAGATGACCTTCTTACC	258
OY	81	HisArgLysAsnTyrValArgValValSerArgAsnGlyGluSerSerGluLeuAspLeu	100
Db	259	CACCGCAAGACCTACGTACAGCGGTGTGAGCCGGAATGAGAACTCAGTAACCTGACTTA	318
OY	101	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle	120
Db	319	CAGGGCATCCCAAAATGGACTCATATATACGGGCAACCTCAAGTGTGGCTGTGAGAGCATT	378
OY	121	ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys	140
Db	379	GTGAGAGAAATCGAGGATGACCTTATGAAATTTCTTCTCCACAGAGGGGTGACAAAGTTAAA	438
OY	141	AspLysLeuGlySerLysArgThrAspLeuGlyAspHisAlaLeuHisIleArgSerHisAsp	160
Db	439	GACAAACCTTTCAGCTAAGCGGACAGATCTATGTACCATGCGCCTCGACAGATCTCAGAT	498
OY	161	GluLeu 162	

Db	99	GAGCTA	504	RESULT_15
LOCUS	BM832828			
DEFINITION	BM832828	588 bp	mRNA	linear
ACCESSION	K-EST0107375	SS5NU484s1	Homo sapiens	CDNA clone SS5NU484s1-11-B03
VERSION	5,			mRNA sequence.
KEYWORDS	BM832828			
SOURCE	BM832828.1	GI:19189237		
ORGANISM	human.			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
	1 (bases 1 to 588)			
	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.			
	21C Frontier Korean EST Project 2001			
	Unpublished (2002)			
	Contact: Kim YS			
	Genome Research Center			
	Korea Research Institute of Bioscience & Biotechnology			
	52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea			
	Tel: +82-42-860-4470			
	Fax: +82-42-860-4409			
	Email: yongsung@mail.kribd.re.kr			
	Plate: 11 row: B column: 03			
	High quality sequence stop: 588.			
	Location/Qualifiers			
	1..588			

BASE COUNT  
ORIGIN

149 a 140 c 176 g 123 t

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="SSSNU484s1-11-B03"  
 /clone\_11b="SSSNU484s1"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_type="Epithelial"  
 /cell\_line="SMU-484"  
 /lab\_host="Top10F"  
 /note="Organ: Stomach; Vector: pTZ189P1; Site\_1: EcoRI;  
 Site\_2: NotI; The poly (A)+ RNA was decapped with tobacco  
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
 including EcoRI site by treatment of T4 RNA ligase. The  
 first strand cDNA was synthesized from oligo dt-selected  
 mRNA by priming with dr-tailed vector. The dr-tailed  
 vector was adjusted to have about 60nt. The cDNA vector  
 was circularized with E. coli DNA ligase after digestion  
 of EcoRI which site is also included in vector. An RNA  
 strand converted to a DNA strand by Okayama-Berg method.  
 The obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10F by electroporation method.  
 After analyzing and sequencing about 2,000 - 3,000  
 colonies in original cDNA library, the abundant cDNAs were  
 selected and amplified by PCR reaction using vector region  
 primer including T7 promoter as 5' primer and N(drr)14 as  
 3' primer. The PCR products were used as template for  
 synthesis of biotinylated single stranded RNA by in vitro  
 transcription reaction. The synthesized RNA probes were  
 hybridized with antisense single stranded cDNAs prepared  
 from original library and incubated with avidin-gel.  
 After removing DNA-RNA hybrids by centrifuge, the  
 subtracted cDNA libraries were constructed by  
 transformation of the remaining DNA into competent cells E.  
 coli Top10F with electroporation method."

BASE COUNT	149	a	140	c	176	g	123	t
ORIGIN								
Alignment Scores:								
Pred. No.:	1,26e-97		Length:		588			
Score:	830.00		Matches:		156			
Percent Similarity:	98.77%		Conservative:		2			
Best Local Similarity:	97.53%		Mismatches:		2			

Db 258 CGAAGAGCCAGATCTACACTGTGAGGCTCTGCTGTGATGATTAAGACTGC 317  
 QY 21 GtUilealArValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 40  
 Db 318 GAAATTCGCCGCTGGACCCCAAGAACCATTCAGATGGATCTTCCGATCATCA 377  
 QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
 Db 378 GATGGCAGCAGTACGATGGAGGTACTTATAGCCCGCTCAGAGGCCACCTCACAGAG 437  
 QY 61 LeuLeuGlnGluValCysAspArgMetGlyGluTyrGlyGlnIleAspProSerThr 80  
 Db 438 TTCTCTGAGAGGTGTCTGTGACCAATGAAGAGATACGGGAAACAGATTCCTTAC 497  
 QY 81 HisArgLysAntTyrValArgValValSerArgAsnGlyLysSerGlnLeuAspLeu 100  
 Db 498 CACCGCAAGAACTACGATACCGCTGTGAGCGGAAATGGAATCCAGTGAACCTGACTTA 557  
 QY 101 GlnGlyTyrLeuArgIleAspSerAspIleSerGlyThrLeuLysPheAla-CysGlnSerIle 120  
 Db 558 CAGGCGATCCGAATTCAGATATCAGATACGCGACCCCTCAAGTTTGGCTGTGAGACAT 617  
 QY 120 eValGlnGluTyrGlnAspGlnLeuIleGlnPheSerArgGlnAlaAspAsnValLy 140  
 Db 618 TGTGGAAGAAATACGAGATAGCTTATCGAATTCCTCTCCAGAGAGGCTACAACTTAA 677  
 QY 140 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAs 160  
 Db 678 AGCAAAACTTGGCAGTAAAGGAGACAGATCTATGTGACCAATGCCCTGCACAGATCTCAGA 737  
 QY 160 pGluLeu 162  
 Db 738 TGAACCT 744  
 RESULT 13  
 B1411532  
 LOCUS B1411532 709 bp mRNA linear EST 14-AUG-2001  
 DEFINITION 602965126F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5120574.5',  
 mRNA sequence.  
 ACCESSION B1411532  
 VERSION B1411532.1 GI:15172455  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NIH-MGC http://mgi.nci.nih.gov/  
 1 (bases 1 to 709)  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
 http://image.llnl.gov  
 plate: LLAM11294 row: c column: 07  
 high quality sequence start: 4  
 high quality sequence stop: 701.  
 Location/Qualifiers  
 1..709  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone\_image="5120574"  
 /clone\_lib="NCI\_CGAP\_Lu33"  
 /issue\_type="pooled lung tumors"  
 /lab\_host="DH10b (phage-resistant)"  
 /note="Organ: Lung; Vector: pT73D-Pac (Pharmacia) with a

modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st  
 strand cDNA was prepared from mRNA obtained from pooled  
 lung tumors with a Not I - oligo(dT) primer [5',  
 TGTACCAATCTGAAGTGGAGAGGCGGCTGTCTTTTCTTTTCTTTT 3'].  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not  
 I and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo. \*  
 BASE COUNT 191 a 170 c 214 g 132 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 6,76e-98 Length: 709  
 Score: 833.00 Matches: 159  
 Percent Similarity: 98.15% Conservative: 0  
 Best Local Similarity: 98.15% Mismatches: 3  
 Query Match: 98.35% Indels: 0  
 DB: 13 Gaps: 0  
 US-10-082-502-21 (1-162) x B1411532 (1-709)  
 QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGlnLeuGluTyr 20  
 Db 164 CGAAGAGCCAGATCTACACTGTGAGGCTCTGCTGTGATGATTAAGACTGC 223  
 QY 21 GtUilealArValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 40  
 Db 224 GAAATTCGCCGCTGGACCCCAAGAACCATTCAGATGGATCTTCCGATCATCA 263  
 QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
 Db 284 GATGGCAGCAGTACGATGGAGGTACTTATAGCCCGCTCAGAGGCCACCTCACAGAG 343  
 QY 61 LeuLeuGlnGluValCysAspArgMetGlyGluTyrGlyGlnIleAspProSerThr 80  
 Db 344 TTCTCTGAGAGGTGTCTGTGACCAATGAAGAGATACGGGAAACAGATTCCTTAC 403  
 QY 81 HisArgLysAntTyrValArgValValSerArgAsnGlyLysSerGlnLeuAspLeu 100  
 Db 404 CACCGCAAGAACTACGATACCGCTGTGAGCGGAAATGGAATCCAGTGAACCTGACTTA 463  
 QY 101 GlnGlyTyrLeuArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlnSerIle 120  
 Db 464 CAGGCGATCCGAATTCAGATATCAGATACGCGACCCCTCAAGTTTGGCTGTGAGACAT 523  
 QY 121 ValGlnGluTyrGlnAspGlnLeuIleGlnPheSerArgGlnAlaAspAsnValLy 140  
 Db 524 GTGGAAGAAATACGAGATAGCTTATCGAATTCCTCTCCAGAGAGGCTGCACAACTTAA 563  
 QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
 Db 584 GACAAATTTGGCAGTAAAGGAGACAGATCTATGTGACCATGCTCCCTGCACAGATCTCAGAT 643  
 QY 161 GluLeu 162  
 Db 644 GAGCTA 649  
 RESULT 14  
 B198865  
 LOCUS B198865 549 bp mRNA linear EST 26-JUN-2000  
 DEFINITION ug83d11.y1 Soares\_mammary\_gland.NMIMG Mus musculus cDNA clone  
 IMAGE:1548981 5' similar to WP.F01F1.11 CE01224 ;, mRNA sequence.  
 ACCESSION B198865  
 VERSION B198865.1 GI:8711034  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

KEYWORDS	EST.
SOURCE	
ORGANISM	house mouse.
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE	1 (bases 1 to 678)
JOURNAL	NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:gspebs@email.nih.gov">gspebs@email.nih.gov</a> Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNM) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNM at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM8523 row: e column: 01 High quality sequence stop: 590.
FEATURES	
source	1..678 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:3486696" /clone_1lb="NCLCGAP_Mam5" /tissue_type="tumor, gross tissue" /dev_stage="7 months" /lab_note="DH10B" /note="Organ: mammary; Vector: pCMV-Sport6; Site:1: SalI; Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"
BASE COUNT	192 a 154 c 191 g 141 t
ORIGIN	
Alignment Scores:	
Pred. No.:	4.69e-98 Length: 678
Score:	834.00 Matches: 162
Percent Similarity:	99.39% Conservative: 0
Best Local Similarity:	99.39% Mismatches: 0
Query Match:	99.47% Indels: 1
DB:	10 Gaps: 0
US-10-082-502-21 (1-162) x BE309953 (1-678)	
OY	1 ArgatserglnaspIeuHiscysgslayalacysargalaleuValaspIleuengIutrp 20 
Db	33 CGAAGGAGCCAGATCTACACTGTGGAGCTTGAGGGCTCTGTGGATGAATTTAGAGTGG 92
OY	21 GluIleuIaaArgValaspProIyLysTrhIleGImetGlySerPheArgIleasnPro 40 
Db	93 GAATTCGCCCGCGGGAGCCCAAGAGACATTCAAGATGGATCTTCGGAATCATCA 152
OY	41 AspIySerGlnSerValValGluValProIyrrAlaArgSerGluAlaHisIeuThrglu 60 
Db	153 GATGGCAGCCAGTCAGTTGGAGGCTTATGCCCCGCTCAGAGCCACCCTCACAGAG 212
OY	61 LeuIeuGluGluValCysAspArgMetLysGluTyrrGluGluIleIaspProSerThr 80 
Db	213 TTGCTTAGAGAGGTGTGTGACCGAATGAAGAGATGAGGGGACAGATTGACCTTTTACC 272
OY	81 HisArgLysasnTyrrValaArgValValSerArgasnGlyGluSerSerGluIeuaspIeu 100 
Db	273 CACCGCAAGAATCTACGTACCGCTGCGAGCCGGAATGGAGATTCAGTGAACCTAGACTTA 332
OY	101 GlnGlyIleArgIleaspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120 
Db	333 CAGGGCATCCGAATTTGATCATGATATACAGCGGAGACCTCAAGTTGGCTGTGAGAGCATTT 392
OY	121 ValGluGluTyrrGluaspGluIeuIleGluIlePhePheSerArgGluAlaaspasnVallys 140

Db	393	GTCGAAGAATACGAGAGTAGACTTATCCAAATCTCTCCAGAGAGCGTGACAAGCTTAA	452
Oy	141	AsplysiLeucYsser-LysArgThrAspleuCysaspHisAlaLeuHisArgSerHisAs	160
Db	453	GACAAACCTTGCAGTMAAGCGCAGCATCTATGTACCATGCCTCCGACAGATCTCACGA	512
Oy	160	peluteud 162	
Db	513	TGAGCTA 519	
RESULT 12			
LOCUS	B1408381	891 bp	mRNA linear EST 14-AUG-2001
DEFINITION	602964013F1 NCL_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5119563 5'		
ACCESSION	B1408381		
VERSION	B1408381.1		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 891) NIH-MGC http://mgs.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC)		
AUTHORS	Unpublished (1999)		
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.		
COMMENT	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM11291 row: 1 column: 04 High quality sequence stop: 873. Location/Qualifiers 1..891 /organism="Mus musculus" /strain="CZECH IT" /db_xref="taxon:10090" /clone_image="IMAGE:5119563" /clone_id="NCL_CGAP_Lu33" /tissue_type="pooled lung tumors" /lab_host="DH10B (phage-resistant)" /note="Organ: Lung; Vector: pRT7D-Pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5' TGTTCACATCTGCATGAGTGGAGCGCGCTCTGTCTTTTCTTTTTTTT 3'] . Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."		
FEATURES			
source			
BASE COUNT	240 a	215 c	258 g 178 t
ORIGIN			
Alignment Scores:			
Pred. No.:	7e-98	Length:	891
Score:	834.00	Matches:	162
Percent Similarity:	99.39%	Conservative:	0
Best local Similarity:	99.39%	Mismatches:	0
Query Match:	98.47%	Indels:	1
DB:	13	Gaps:	0
US-10-082-502-21 (1-162) x B1408381 (1-891)			
Oy	1	ArgArgSerGlnAspPheUHisCysGlyAlaCysArgAlaLeuValAspGluDeuGluTrp	20

Pred. No.: 3,83e-98 Length: 724  
 Score: 835.00 Matches: 159  
 Percent Similarity: 98.77% Conservative: 1  
 Best Local Similarity: 98.15% Mismatches: 2  
 Query Match: 98.58% Indels: 0  
 DB: 14 Gaps: 0

US-10-082-502-21 (1-162) x BQ210416 (1-724)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 |||||  
 DB 629 CGAAGAGCAGCAGATCTACACTGTGAGGCTTGAGGGCTGTGTGATGATTAAGAGTGG 570  
 |||||  
 QY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetCysSerPheArgIleAsnPro 40  
 |||||  
 DB 569 GAAATTCGCCCGCTGGACCCCAAGAACCATTCAGATGGGATCTCCGAATCAATCCA 510  
 |||||  
 QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 |||||  
 DB 509 GATGGCAGCAGTCACTGAGGAGTTCCTTATGCCCGCTCAAGAGCCCTCACGGAG 450  
 |||||  
 QY 61 LeuLeuGluGluValAlaCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 |||||  
 DB 449 TTGCTTGAGAGAGATATGTGACCGAATGAAGAGATATGGGAAACAGATGACCTTCTACC 390  
 |||||  
 QY 81 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeu 100  
 |||||  
 DB 389 CACCGCAAGAACTACGTCCTGTGTCGGCCGAGACGAGATCCAGTAACATAAGACTTA 330  
 |||||  
 QY 101 GlnGlyTyrLeuArgTyrLeuAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 |||||  
 DB 329 CAGGGATATCCGAATCGAATTCAGATATCAGAGGACCTCAAGTTGGGTGAGAGCATTT 270  
 |||||  
 QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
 |||||  
 DB 269 GTGAGAGAGTACAGAGATGAATTAATTAATTTCTTTCAAGAGAGCTGCAACGTTAAA 210  
 |||||  
 QY 141 AspLysLeuLysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
 |||||  
 DB 209 GACAACTTTGACAGTAAGCGAGACAGATCTATGTAGACCATCCCTGCACAGATCTCATGAT 150  
 |||||  
 QY 161 GluLeu 162  
 |||||  
 DB 149 GAGCTA 144  
 |||||  
 RESULT 10 802 bp mRNA linear EST 01-APR-2001  
 LOCUS BF608664  
 DEFINITION MYL\_00106 Mouse 9-day fetus cDNA library ICRFp522 Mus musculus  
 accession BF608664  
 version BF608664.1 GI:13505156  
 keywords EST.  
 source house mouse.  
 organism Mus musculus.  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 802)  
 Yahyawi, M., Hennig, S., Neidhardt, L., Radelof, U., Hermann, B.G.,  
 Lehman, H. and O'Brien, J.  
 Detection of a high number of novel genes in a 9-day mouse embryo  
 cDNA library normalised by oligonucleotide fingerprinting  
 unpublished (2001)  
 JOURNAL  
 COMMENT  
 Contact: Hennig S  
 Laboratory 123, dept. Lehrach  
 Max-Planck-Institut fuer Molekulare Genetik  
 Ihnestr. 63-73, D-14195 Berlin, Germany  
 Tel: +49 30 8413 1612  
 Fax: +49 30 8413 1380  
 Email: hennig@molgen.mpg.de  
 EST's are made from clones being representatives of clone clusters.  
 Clone clusters were calculated from oligonucleotide fingerprints.  
 PCR Primers

FEATURES  
 source  
 1. 802  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="ICRFp522B2440"  
 /clone.lib="Mouse 9-day fetus cDNA library ICRFp522"  
 /tissue\_type="whole embryo"  
 /dev\_stage="embryonic 9-day"  
 /lab\_host="E.coli, XL1 blue"  
 /note="Vector: pSVSPORT1; Site.1: NotI; Site.2: SalI;  
 library preparation by oligo-dT priming of RNA. Clones can  
 be ordered from the Resource Center in Berlin,  
 http://www.rzpd.de."

BASE COUNT 193 a 206 c 243 g 159 t 1 others  
 ORIGIN

#### Alignment Scores:

Pred. No.: 4.45e-98 Length: 802  
 Score: 835.00 Matches: 160  
 Percent Similarity: 99.38% Conservative: 1  
 Best Local Similarity: 98.77% Mismatches: 1  
 Query Match: 98.58% Indels: 0  
 DB: 12 Gaps: 0

US-10-082-502-21 (1-162) x BF608664 (1-802)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 |||||  
 DB 300 CGAAGAGCAGCAGATCTACACTGTGAGGCTTGAGGGCTGTGTGATGATTAAGAGTGG 359  
 |||||  
 QY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetCysSerPheArgIleAsnPro 40  
 |||||  
 DB 360 GAAATTCGCCCGCTGGACCCCAAGAACCATTCAGATGGGATCTCCGAATCAATCCA 419  
 |||||  
 QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 |||||  
 DB 420 GATGGCAGCAGTCACTGAGGAGTTCCTTATGCCCGCTCAAGAGCCCTCACGGAG 479  
 |||||  
 QY 61 LeuLeuGluGluValAlaCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 |||||  
 DB 480 TTGCTTGAGAGAGATATGTGACCGAATGAAGAGATATGGGAAACAGATGACCTTCTACC 539  
 |||||  
 QY 81 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeu 100  
 |||||  
 DB 540 CACCGCAAGAACTACGTCCTGTGAGCGGAGAGGAGAAATCCAGTGAATCAATTA 599  
 |||||  
 QY 101 GlnGlyTyrLeuArgTyrLeuAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 |||||  
 DB 600 CAGGGATATCCGAATTCAGATATCAGAGGACCTCAAGTTGGGTGAGAGCATTT 659  
 |||||  
 QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
 |||||  
 DB 660 GTGAGAGAGTACAGAGATGAATTAATTAATTTCTTTCAAGAGAGCTGCAACGTTAAA 719  
 |||||  
 QY 141 AspLysLeuLysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
 |||||  
 DB 720 GACAACTTTGACAGTAAGCGAGACAGATCTATGTAGACCATCCCTGCACAGATCTCAGAT 779  
 |||||  
 QY 161 GluLeu 162  
 |||||  
 DB 780 GAACCT 785  
 |||||  
 RESULT 11  
 LOCUS BE309953  
 DEFINITION 601091903F1 NCI\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:3486596 5',  
 accession BE309953  
 version BE309953.1 GI:9168932

and the oligo-dT track served to verify it as a clone from the  
normalized duodenum library cDNA Library Preparation: M.B. Soares  
Lab Clone distribution: clones will be available through Research  
Genetics (www.resgen.com)  
Seq primer: M13 Forward  
POLYA-Yes.

## FEATURES

SOURCE

Location/Qualifiers

1..674

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="UI-R-EB1-clone-b-02-0-UI"

/clone\_lib="UI-R-EB1"

/dev\_stage="adult"

/lab\_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site 1: Not I; Site 2: Eco RI; UI-R-EB1 is a  
non-normalized Rat Duodenum Library (RDL) constructed in  
pT737 PAC vector according to the procedure described by  
Bonaldi, Lennon & Soares (Genome Research 6:  
791-806, 1996). The oligonucleotide used to prime first  
strand synthesis contained the sequence tag TGTGCTCAT  
between the Not I cloning site and dT18 stretch. The Rat  
Duodenum tissue was provided by Tom Freeman of the Sanger  
Center.

TAG\_LIB=UI-R-EB1

TAG\_TISSUE=duodenum

TAG\_SEQ=TGTGCTCAT

BASE COUNT 143 a 183 c 145 g 203 t  
ORIGIN

## Alignment Scores:

Pred. NO.: 3.45e-98 Length: 674  
Score: 835.00 Matches: 159  
Percent Similarity: 98.77% Conservative: 1  
Best Local Similarity: 98.15% Mismatches: 2  
Query Match: 98.58% Indels: 0  
DB: 14 Gaps: 0

US-10-082-502-21 (1-162) x BQ202181 (1-674)

0Y 1 ArgATGSeRGInAAsPLeuHICyAGValACyARgAlAlaLeuValAspGLuLengLutrp 20  
|||||  
Db 636 CGAAGGACCCAGATCTCACTGAGACCTTGAGGCTGTGGATGAATTAAGTGG 577  
0Y 21 GluIleAlaArGValAsPProLySThrIleGImetGlySerPheArGLIleAsnPro 40  
|||||  
Db 576 GAAATTCGCCGCGGTGAGCCCAAGAACATTCAGATGGGATCCTTCGAATCATCA 517  
0Y 41 AspGLySerGInSerValAlaGLuValProGyralaArGSeRGluAlaHISLeuThrGlu 60  
|||||  
Db 516 GATGCGACCCAGTCAGTTGTGAGAGTTCCTTAATGCCCTCAGAGGCCACCTCAGGAG 457  
0Y 61 LeuIleuGLuGLuValCyAsPArGmetLySGluGyrgluGluInIleAsPProSerThr 80  
|||||  
Db 456 TTGCTTGGAGATATGTGACCGAATGAAGATGTGGGAACAATGACCTTCTAC 397  
0Y 81 HisArGLyAsnTYrValArGValAlaSerArGAnGLyGluSerSeRGluLeuAsPLeu 100  
|||||  
Db 396 CACGCGAAGAACTACGTCGTTGTGCGCGGAGAACGAGATCCAGTAAGTACTTA 337  
0Y 101 GlnGLyIleArGLIleAsPSeRAsPLeSerGLyThrLeuLysPheAlaCySGluSerIle 120  
|||||  
Db 336 CAGGTTATCCCAATTCAGATTCAGATTCAGTGCACCCCTCAAGTTGCTGTAGAGCAT 277  
0Y 121 ValGLuGLuTYrGLuAsPGLuLeuIleGluPheSerArGValAlaAsPAsnValLys 140  
|||||  
Db 276 GTGGAGGATGAGAGATGAATTAATTAATCTTTTAAAGAGGCTGACAAACGTTAAA 217  
0Y 141 AspLyLeuCySerLyArGTThrAsPLeuCyAsPHisAlaLeuHISArGSeRHisAsP 160  
|||||  
Db 216 GACAAACTTTCAGTAAACGAGACAGATCTATGTGACATGCCCTGCACAGATCTCATAT 157

0Y 161 GluLeu 162  
|||||  
Db 156 GAGCTA 151

## RESULT 9

BQ210416/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

1 (bases 1 to 724)  
Bonaldi,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. The sequence tag present in the cDNA between the NotI site  
and the oligo-dT track served to verify it as a clone from the  
normalized cartilage library cDNA Library Preparation: M.B. Soares  
Lab Clone distribution: clones will be available through Research  
Genetics (www.resgen.com)  
Seq primer: M13 Forward  
POLYA-Yes.

## FEATURES

SOURCE

Location/Qualifiers

1..724

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="UI-R-DY1-clone-e-06-0-UI"

/clone\_lib="UI-R-DY1"

/tissue\_type="Cartilage"

/dev\_stage="Adult"

/lab\_host="DH10B (Life Technologies)"

/note="Organ: Femur and Tibia; Vector: pT73D-Pac  
(Pharmacia) with a modified polylinker. Site 1: Not I;  
Site 2: Eco RI; UI-R-DY1 is a normalized cDNA library  
containing the following tissue(s): Rat Cartilage from  
Femur and Tibia. The library was constructed according to  
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,  
1996. First strand cDNA synthesis was primed with an  
oligo-dT primer containing a Not I site. Double stranded  
cDNA was ligated to an EcoR I adaptor, digested with Not  
I, and cloned directionally into pT73-Pac vector. The  
oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tag for this library is CTATGAGCG. The Rat  
cartilage tissue was provided by Dr Jeff Stevens at the  
University of Iowa.

TAG\_LIB=UI-R-DY1

TAG\_TISSUE=cartilage

TAG\_SEQ=CTATGAGCG

BASE COUNT 156 a 206 c 157 g 205 t

ORIGIN

Alignment Scores:

Db	395	CATGACAGACAGACTAGTTGTGTGGAGGTACCTTATATGCCCGCTGAGAGCCACACTCCACAGAG	454
Qy	61	LeuLeuGluGluValCysAspArgMetIysGluTyrGluGluGlnIleAspProSerThr	80
Db	455	TTGCTTTGAGGAGAGTGTGTGACCGAATGAAGAGTCCGGGGAAACAATTTGACCCCTTCTACC	514
Qy	81	HisArgLysAsnTyrValArgValValSerArgAsnGluGluSerSerGluLeuAspLeu	100
Db	515	CACGCCAAGACTACGTACGCTGCGCTGAGCCGGAAATGGAGAAATCCAGTGAACACTTA	574
Qy	101	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle	120
Db	575	CAGGCGATCCCGAATTTGACTCAGATATTCAGCGGCACCCCTCAAGTTTGCCTGTGACAGCAAT	634
Qy	121	ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys	140
Db	635	GTCGAGAAATACGAGAGATGAGCTTATCCGAATCTTCTCCAGAGAGCGGTACAACTTAAA	694
Qy	141	AspLysLeuCysSerLysAlaThrAspLeuCysAspAsnIleAlaLeuAlaLysSerHisAsp	160
Db	695	GACAAACTTTGCAGTAAAGCGACAGATATGATGTGACASCATGCCCTCAGAGATCTCAGCAT	754
Qy	161	GluLeu 162	
Db	755	GAGCTA 760	

RESULT 7	LOCUS	DEFINITION	ACCESSION
AW919569/c	AW919569	665 bp mRNA linear EST 25-MAY-2000	ESR350873 Rat gene index, normalized rat, norvegicus, Bento Soares
			Rattus norvegicus cdna clone RGIF21 5' end, mRNA sequence.
			AW010560

Accession	AM13509	GI:8085366
Version	AW19569.1	
Keywords	EST.	
Source	Norway rat.	
Organism	Rattus norvegicus	

**REFERENCE**  
**AUTHORS**

Euteleostomi; Euteleostomi;  
Euteleostomi; Euteleostomi;  
Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 665)  
Lee, N.H., Glodet, A., Chandra, I., Mason, T.M., Quackenbush, J.,  
Kerlavage, A.R., and Adams, M.D.

TITLE  
Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat  
Gene Index  
JOURNAL  
Unpublished (1998)  
COMMENT  
Contact: Lee, NH

The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: [nhlee@tigr.org](mailto:nhlee@tigr.org)  
This clone is available through the ATCC, contact the ATCC  
tel#703-365-2700 for further information  
Seq primer: M13 Reverse.

FEATURES	Location/Qualifiers
source	1. .665

BASE COUNT	ORIGIN
150 a	186 c 148 g 161 t

Alignment Scores:	
Pred. No.:	3,38e-98
Score:	835.00
Length:	665
Matches:	159

Percent Similarity:	98.77%	Conservative:	1
Best Local Similarity:	98.15%	Mismatches:	2
Query Match:	98.58%	Indels:	0
DB:	10	Gaps:	0

US-10-082-502-21 (1-162) x AW919569 (1-665)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20

Db 620 CGAAGAGCCAGACTTACACTGTGAGCTTGCAGGCTCTGGTGGATGATTAGAGTGG 56

Qy 21 GIULIETARGVALASPROLYSLYSTRHILEGIMETGLYCERPHEARGILEASNPRO 40  
|||||

Db 560 GAATTTGCCCCGCTGACCCCAAGAGACCATTCAGATGGATCCTCCGATCAATCCA 50

41 AspglySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60

6] [auIauG]uc[ uVa][CvsaasaaarMetIwsc[uTtwrc]wc[uc]nrlaaasndroSae-ahr 80

440 TTGCTGGAGATATGTGACCAATGAAGAGATAGGGGACAGATTGACCTTTACC 38

81 HISARGLYSASNFTYRVALARGVALVALSERARGASNGLYGLUSERSERGLULEUASPLeu 100

Db  
380 CACCGCAGACTACGTCGGTGTGTGCGCCGGAACGAGATTCAGTACCTAGACTTA 322

101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 122

Db 320 CAGGGTATCCGAATCGATTACAGATATCAGTGGCACCCTCAAGTTTGGCTGTGAGAGCAAT 26

QY 121 VALGLUGLUTYRGLUASPGLULEULEGLUPHESPESERRARGGLUALAASPAASVALYS 14

Db 260 GTGGAGAGTACGAGATGAACCTATTGAACTCTTTCACAGAGAGGCTGACCAACGTTAA 20

QY 141 AsplysleucysSerIysArgThrAspLeucysAspHisAlaLeuHisArgSerHisAsp 166

Db 200 GACAAACTTTGCAGTAAAGCGACAGATCTATTGTGACCATTGCCCTGCACAGATCTCATGAT 14

qy	161	Guileu	162
Db	140	caacca	135

RESULT A

DB	140	GRACIA	123
----	-----	--------	-----

BQ202181/c	
Locus	
BQ202181	
674 bp	
mRNA	linear
EST	02-MA

DEFINITION	
UI-R-EB1-cle-b-02-0-UI.s1	UI-R-EB1 Rattus norvegicus cDNA clone
UI-R-EB1-cle-b-02-0-UI.3'	mRNA sequence.

ACCESSION	BQ202181
VERSION	BQ202181.1
	GI:20418646

KEYWORDS	EST.
SOURCE	Norway rat.
ORGANISM	Dattus norvegicus

ORGANISM  
natus holveticus  
Eukaryota; Metazoa;  
Mammalia; Eutheria;  
Rodentia; Sciuro-  
mammalia; Muridae;  
Murinae

REFERENCE  
1 (bases 1 to 674)

AUTHORS	Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE	Normalization and subtraction: two approaches to facilitate q

discovery  
Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477  
Contact: Soares, MB  
Address: 604 West Elmwood

Program for rat gene discovery and mapping  
University of Iowa  
[Research Building Iowa City, IA 52242  
451 Eastlain Medical]

421 Eastern Medical Research Building Iowa City, IA 52242, U.S.A.  
Tel: 319 335 8250  
Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu  
The sequence contained an oligo-dT track that was present in

oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI s







OY 161 Glutenu 162  
 DB 659 GACCTT 664

RESULT 4  
 BI412539 843 bp mRNA linear EST 14-AUG-2001  
 LOCUS BI412539  
 DEFINITION 60299095F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5146967 5',  
 mRNA sequence.  
 BI412539  
 ACCESSION BI412539  
 VERSION BI412539.1 GI:15173462  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 843)  
 REFERENCE NIH-MGC http://mgc.ncl.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaaps-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LAM11362 Row: n Column: 24  
 High quality sequence start: 38  
 Location/Qualifiers  
 1..843  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5146967"  
 /clone\_11b="NCI\_CGAP\_Lu33"  
 /tissue\_type="pooled lung tumors"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: Lung; Vector: pTZ19-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: NotI; Site\_2: EcoRI; 1st  
 strand cDNA was prepared from mRNA obtained from pooled  
 lung tumors with a Not I - oligo(dT) primer [5',  
 TGTTCACATCTGAGAGGAGGCGGCGCTCTTTTCTTTTCTT 3'].  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not  
 I and Eco RI sites of the modified pTZ19 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 239 a 195 c 238 g 171 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,31e-99 Length: 843  
 Score: 847.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-10-082-502-21 (1-162) x BI412539 (1-843)

OY 1 ArgArgSerGlnAspLeuHIsGyGlyAlaCysArgAlaLeuValAspGluLeuGluTyr 20  
 DB 214 CGAAGGAGCCCAAGATCTTACACTGTGAGCTTGCGAGGCTCTGGTGATGATTTGAGAGTGG 273

OY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAspPro 40  
 DB 274 GAAATTCCCGCGTGGAGCCCAAGAGACCATTCAGATGGATCTTCCGATCAATCA 333

OY 41 AspGlySerGlnSerValGluValProTyrAlaArgSerGluAlaIleLeuThrGlu 60  
 DB 334 GATGGACAGCCAGCAGTGTGTGAGAGTACTTATGCGCCCTCAGAGGCCACCTGCAGACAG 393

OY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 DB 394 TTGCTGTAGGAGGTGTGTGACCCGAATGAAAGAGTACGGGGAAACAGATTGACCTTCTACC 453

OY 81 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyGlnSerSerGluLeuAspLeu 100  
 DB 454 CACCGCAAGAACTACGTACGGCTGTGACCGCGAATGAGAAATCCAGTAAGTACACTTA 513

OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 DB 514 CAGGCAATCCGAATTTGACTCAGATATCAGCGCACCCCTCAAGTTTGGCTGTGAGAGCAT 573

OY 121 ValGluGluTyrGluAspGluLeuIleLeuPheSerArgGluAlaAspAsnValLys 140  
 DB 574 GTGGAAAGAAATCGAGGATGACCTTATCGAATCTTCTCCAGAGAGGCTGACACAGTTTAA 633

OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHIsArgSerHisAsp 160  
 DB 634 GACAACTTTCAGATGAGCGACAGATATATGACCATGCCCTTCACAGATCTCAGCAT 693

OY 161 Glutenu 162  
 DB 694 GACCTA 699

RESULT 5  
 AK013014 853 bp mRNA linear HTC 19-JAN-2002  
 LOCUS AK013014  
 DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length  
 enriched library, clone:2810406124:transmembrane protein 4, full  
 insert sequence.

ACCESSION AK013014  
 VERSION AK013014.1 GI:12850114  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (strain: C57BL/6J) 10, 11 days embryo cDNA to mRNA,  
 clone\_11b:RIKEN full-length enriched mouse cDNA library  
 clone:2810406124.

ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE Carninci, P. and Hayashizaki, Y.  
 1 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636

REFERENCE Itoh, M., Kono, H., Okazaki, Y., Sugahara, Y., Shibata, K.,  
 2 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159

REFERENCE Shibata, K., Itoh, M., Aizawa, K., Nagaoke, S., Sasaki, N., Carninci, P.,  
 3 Kono, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multichannel sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861

REFERENCE Kawai, J., Shitagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

/db\_xref="MCD:MGI:1928477"  
 /translation="MKGMWLLALLGVLLGTAMARRSODLHCACRALVDELEWEIAR  
 VPKRTIQMSRINDPSQSYVEVPYARSEAHTELEEVCDRMKEVEQIDPSTHR  
 KKYVRVRSNGSESLDLOGIRIDSDISCTLKFACSEIVEEDELIEFFSHEDNVK  
 DKCSKRPMDLGDHALHRSDEL"

BASE COUNT 210 a 180 c 225 g 157 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.15e-99 Length: 772  
 Score: 847.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 11 Gaps: 0

US-10-082-502-21 (1-162) x AK007914 (1-772)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 |||||  
 DB 164 CGAAGAGCCGCAAGATCTACACTGTGGAGCTTCGAGGCGCTCGTGGAATTAGAGTGG 223  
 QY 21 GlnIleAlaArgValAspProLysLysThrIleGlnMetCysSerPheArgIleAsnPro 40  
 |||||  
 DB 224 GAAATTCGCCGCTGGACCCCAAGAGACCATTCAGATGGGATCTTCCGAATCAATCCA 283  
 QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
 |||||  
 DB 284 GATGGCAGCCAGTCAGTGTGGAGTACCTTATGCCCGCTCAGAGGCCACCTCACAGAG 343  
 QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGlnIleAsnProSerThr 80  
 |||||  
 DB 344 TTCTCTGAGAGGCTGTGTGACCAATGAGAGTACGCGGAGACAGATTCACCTTCACCC 403  
 QY 81 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyLysSerGluLeuAspLeu 100  
 |||||  
 DB 404 CACCGCAAGACTACGACCTGCTGAGCCGGAATGGAATTCAGTACTAGACTTA 463  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 |||||  
 DB 464 CAGGGCATCCGAATTCAGATTCAGATTCAGGCGCACCTCAAGTTGGCTGTGAGAGCAT 523  
 QY 121 ValGluGluTyrGluAspGluLeuIleGlnPheSerArgGluAlaAspAsnValLys 140  
 |||||  
 DB 524 GTGGAAGAAATACAGAGTATGCTTATTCGAATCTTCTCCAGAGAGGCTGACAACTTAAA 583  
 QY 141 AspLysLeuSerLysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
 |||||  
 DB 584 GACAACTTTCAGATTAAGCGGACAGATCTATGTGACCATCCCTGACACAGATCTCAGAT 643  
 QY 161 GluLeu 162  
 |||||  
 DB 644 GAGCTA 649

RESULT 3  
 BI408525 808 bp mRNA linear EST 14-AUG-2001  
 LOCUS 602964995F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5120471 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI408525  
 VERSION BI408525.1 GI:15169448  
 KEYWORDS EST.  
 ORGANISM house mouse.  
 SOURCE Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 808)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldi, Ph.D.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA sequencing by: Incyte Genomics, Inc.  
 Clone distribution: NCI-CCAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1M11293  
 http://image.llnl.gov  
 High quality sequence start: 4  
 High quality sequence stop: 768.  
 Location/Qualifiers

FEATURES  
 source

1.808  
 /organism="Mus musculus"  
 /strain="CZECH II"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5120471"  
 /clone\_1lb="NCI\_CGAP\_Lu33"  
 /tissue\_type="Pooled lung tumors"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a  
 modified polylinker. Site 1: NotI; Site 2: EcoRI; 1st  
 strand cDNA was prepared from mRNA obtained from pooled  
 lung tumors with a Not I - oligo(dT) primer [5'  
 TCTTACCACTGTGAAGTGGAGCGCGCCCTCTGTTTCTTTTCTTTT 3']  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not  
 I and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldi."

BASE COUNT 224 a 181 c 237 g 166 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.23e-99 Length: 808  
 Score: 847.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-10-082-502-21 (1-162) x BI408525 (1-808)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 |||||  
 DB 179 CGAAGAGCCGCAAGATCTACACTGTGGAGCTTCGAGGCGCTCGTGGAATTAGAGTGG 238  
 QY 21 GlnIleAlaArgValAspProLysLysThrIleGlnMetCysSerPheArgIleAsnPro 40  
 |||||  
 DB 239 GAAATTCGCCGCTGGACCCCAAGAGACCATTCAGATGGGATCTTCCGAATCAATCCA 298  
 QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
 |||||  
 DB 299 GATGGCAGCCAGTCAGTGTGGAGTACCTTATGCCCGCTCAGAGGCCACCTCACAGAG 358  
 QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGlnIleAsnProSerThr 80  
 |||||  
 DB 359 TTCTCTGAGAGGCTGTGTGACCAATGAGAGTACGCGGAGACAGATTCACCTTCACCC 418  
 QY 81 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyLysSerGluLeuAspLeu 100  
 |||||  
 DB 419 CACCGCAAGACTACGACCTGCTGAGCCGGAATGGAATTCAGTACTAGACTTA 478  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 |||||  
 DB 479 CAGGGCATCCGAATTCAGATTCAGATTCAGGCGCACCTCAAGTTGGCTGTGAGAGCAT 538  
 QY 121 ValGluGluTyrGluAspGluLeuIleGlnPheSerArgGluAlaAspAsnValLys 140  
 |||||  
 DB 539 GTGGAAGAAATACAGAGTATGCTTATTCGAATCTTCTCCAGAGAGGCTGACAACTTAAA 598  
 QY 141 AspLysLeuSerLysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
 |||||  
 DB 599 GACAACTTTCAGATTAAGCGGACAGATCTATGTGACCATCCCTGACACAGATCTCAGAT 658

Db	385	CACCCAGAACACTAGCTAGCGCTCGTAGCCGGAATGGAGATCCAGTACACTAGACTTA	444
Oy	101	GINGLYIILEATGIIAASPseraspIleSerGlyTHrLeuLysPheAlaCysGluSerIle	120
Db	445	CAGGCATCCGCAATGACCTACAGATATCAGCCGACCCCTCCMACTTGGCGGTAGAGCAATT	504
Oy	121	VALGLUGIUTYGIUASpGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys	140
Db	505	GTGGAGAAGATACGAGAGATGAGCTTATCCGAATCTCTCCACAGAGAGCTGACACCTTAA	564
Oy	141	AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisasp	160
Db	565	GACCAACTTTCAGTAAAGGAGCAGATGATCTATGTGACCATGCGCTCAGAGATCTCAGCAT	624
Oy	161	GluLeu 162	
Db	625	GAGCTA 630	
RESULT 2			
LOCUS	AK007914		
DEFINITION	AK007914	772 bp	mRNA
LOCUS	AK007914		linear
DEFINITION	Mus musculus 10 day old male pancreas cDNA, RIKEN full-length		
DEFINITION	enriched library, clone:1810060D19:transmembrane protein 4, full		
DEFINITION	insert sequence.		
DEFINITION	AK007914		
DEFINITION	AK007914.1	GI:12841773	
DEFINITION	HTC: CAP trapper.		
DEFINITION	Mus musculus (strain:C57Bl/6J) 10 day old male pancreas cDNA to		
DEFINITION	mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library		
DEFINITION	clone:1810060D19.		
DEFINITION	Mus musculus		
DEFINITION	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
DEFINITION	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
DEFINITION	1		
DEFINITION	Carninci, P. and Hayashizaki, Y.		
DEFINITION	High-efficiency full-length cDNA cloning		
DEFINITION	Meth. Enzymol. 303, 19-44 (1999)		
DEFINITION	99279253		
DEFINITION	10349636		
DEFINITION	2		
DEFINITION	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
DEFINITION	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
DEFINITION	Normalization and subtraction of cap-trapper-selected cDNAs to		
DEFINITION	prepare full-length cDNA libraries for rapid discovery of new genes		
DEFINITION	Genome Res. 10 (10), 1617-1630 (2000)		
DEFINITION	11042159		
DEFINITION	3		
DEFINITION	Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,		
DEFINITION	Kono, H., Akiyama, J., Nishi, K., Kitunai, T., Tashiro, H., Itoh, M.,		
DEFINITION	Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,		
DEFINITION	Yamanoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kasai, K.,		
DEFINITION	Fujiwara, S., Inoue, K., Togawa, Y., Iwata, M., Ohara, E., Watabiki, M.,		
DEFINITION	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,		
DEFINITION	Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.		
DEFINITION	RIKEN integrated sequence analysis (RISA) system--384-format		
DEFINITION	sequencing pipeline with 384 multicapillary sequencer		
DEFINITION	Genome Res. 10 (11), 1757-1771 (2000)		
DEFINITION	20530913		
DEFINITION	11076861		
DEFINITION	4		
DEFINITION	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,		
DEFINITION	Atkawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,		
DEFINITION	Aikawa, K., Iwata, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,		
DEFINITION	Salto, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Salto, R.,		
DEFINITION	Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavola, T.,		
DEFINITION	Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochwa, H.,		
DEFINITION	Kuehl, P., Lewis, S., Matsuo, Y., Niki, K., Ito, T., Suzuki, R., Tomita, M.,		
DEFINITION	Okouchi, S., Schmitt, L. M., Stanbly, F., Suzuki, R., Tomita, M.,		
DEFINITION	Wagner, L., Washio, T., Sakai, K., Okita, T., Furuno, M., Aono, H.,		
DEFINITION	Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,		
DEFINITION	Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bull, C.,		

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 5 (bases 1 to 772)  
 Aachchi,J., Alizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,  
 Atakawa,T., Baldarelli,R., Bono,H., Brownstein,M., But,C.,  
 Carninci,P., Fukuda,S., Fukuishita,Y., Furuno,M., Hanagaki,T.,  
 Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiroaka,T., Horii,F.,  
 Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasekawa,T.,  
 Kato,H., Kawaji,J., Kojima,Y., Konno,H., Kouda,M., Kosu,S.,  
 Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,  
 Numasaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,  
 Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Saeki,D.,  
 Schirml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,  
 Sobabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,  
 Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,  
 Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M., and  
 Hayashizaki,Y.  
 Direct Submission  
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,  
 URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
 Fax:81-45-503-9216)  
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
 further details.  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues. First strand cDNA was primed with a primer  
 [5' GAGAGAAGAGAGATCCCAAGACCTCTTTTTCCTTTTTTNN 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse transcriptase  
 and subsequently enriched for full-length by cap-trapper. cDNA went  
 through one round of subtraction to Rot = 20.0. Second strand cDNA  
 was prepared with the primer adapter of sequence[5'  
 GAGAGAAGAGAGCGGCAGTAATTATCTCAGACTATTAATTATGCCCCCCCCC 3'], cDNA  
 was cleaved with XhoI and SclI. Cloning sites, 5' end: XhoI; 3'  
 end: SclI. Host: SOLR.  
 Location/Qualifiers  
 1..772  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM\_DB:1810060D19"  
 /db\_xref="MGD:MGI:1897650"  
 /db\_xref="taxon:10090"  
 /clone="1810060D19"  
 /sex="male"  
 /tissue\_type="pancreas"  
 /clone\_id="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="10 day old"  
 1..772  
 /gene="Tmem4"  
 104..652  
 /gene="Tmem4"  
 /note="data source:MGD, source key:MGI:1928477,  
 evidence:ISS  
 putative  
 transmembrane protein 4"  
 /codon\_start=1  
 /protein\_id="BAB25346.1"  
 /db\_xref="GI:12841774"

**REFERENCE**

**AUTHORS** Carninci, P. and Hayashizaki, Y.

**TITLE** High-efficiency full-length cDNA cloning

**JOURNAL** Meth. Enzymol. 303, 19-44 (1999)

**PUBMED** 99279253

**REFERENCE** 10349636

**AUTHORS** 2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

**TITLE** Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

**JOURNAL** Genome Res. 10 (10), 1617-1630 (2000)

**PUBMED** 20493974

**REFERENCE** 11042159

**AUTHORS** 3

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasaiyagi, K., Fujikake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.

**TITLE** RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

**JOURNAL** Genome Res. 10 (11), 1757-1774 (2000)

**PUBMED** 20530913

**REFERENCE** 11076661

**AUTHORS** 4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanka, I., Salto, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Salto, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochwa, H., Knehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Stabill, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hotman, M., Hume, D.A., Kamlay, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombers, P., Noridone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakmoto, N., Sasaki, H., Seto, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyokura, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wysshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohlschütter, S. and Hayashizaki, Y.

**TITLE** Functional annotation of a full-length mouse cDNA collection

**JOURNAL** Nature 409 (6821), 685-690 (2001)

**PUBMED** 21085660

**REFERENCE** 11217851

**AUTHORS** 5 (bases 1 to 754)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Balderelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hangaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Horii, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kuwihara, C., Matsuyama, T., Miyazaki, Y., Nishi, K., Nomura, K., Nunazaki, R., Ono, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Salto, H., Salto, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, Y., Tanaka, T., Teijima, Y., Toyota, T., Yamamura, T., Yananaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

**TITLE** Direct Submission

**JOURNAL** Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,

	COMMENT	Fax:81-45-503-9216) Please visit our web site ( <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> ) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGACGATCCAGACCTCTTTTCTTTTTTNN 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGTGCAGTTAATTAAATTAATCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOUK.
	FEATURES	location/Qualifiers 1..754 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="PANTOM,DB:2900022B12" /db_xref="MGD:MGI:1896245" /db_xref="taxon:10090" /clone="2900022B12" /sex="male" /tissue_type="hippocampus" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 1..754 /gene="Ymen4" 85..633 /gene="Ymen4" /note="data source:MGD, source key:MGI:1928477, evidence:ISS putative transmembrane protein 4" /codon_start=1 /protein_id="BAB28909.1" /db_xref="GI:12850980" /db_xref="MGD:MGI:1928477" /translacion="MKMGWLTALLGLVLTGMARRSODLHCACRALVDLEWEIAR VDPKSLTOMGSFRINPDGOSGVVEVPARSBAHLTELEVCDRMKEGBOIDSTSHRR KNRYRVSRNSESSELDLOGRIDSDISGLTKFACESIVEYEDELIEFPSREADNYK DKLSKRTRDLCDHALHRSDBEL"
	gene	
	CDS	
	BASE COUNT	200 a 175 c 220 g 159 t
	ORIGIN	
	Alignment Scores:	
	Pred. No.:	1 11e-99                  Length: 754
	Score:	847.00                  Matches: 162
	Percent Similarity:	100.00%                Conservative: 0
	Best local Similarity:	100.00%                Mismatches: 0
	Query Match:	100.00%                Indels: 0
	DB:	11                        Gaps: 0
OY	US-10-082-502-21 (1-162) x AK013568 (1-754)	
OY	1 ATGATGSeRGinAspLeuNHISCSySLyAlAcYSATgAlALeUvAlAspLUleUGluTrp 20	
Db	145 CGAAGGAGCGCAAGACTCTACACTGTGGAGCTTGACAGCGCTGTGTGAGTGAATTAAGTGG 204	
OY	21 GUUUUAUAtgValAspProLyLysThrllEglmEtGlySerPheatrgllEAnPro 40	
Db	205 GAATAATGCCCGCGTGGAGCCCAAGAAGACATTCAGATGGGATCCTTCGGAATCAATCCA 264	
OY	41 AspgLYserGinSeRValVaIGluValAProtyrAlaArgSerGUAlAHIsLeuhtGlu 60	
Db	265 GATGGCACCGACtCGTGTGGAGGTACTTATGCCCCTCAGAGGCCACCCTCCACAAG 324	
OY	61 LeuLeugLuGlUvAlCyASApArqMetCYsgLUtYrGluGlUGInIlleAspProSerThr 80	
Db	325 TTGGCTTGAAGGAGAGTGTGTGACCGAATGAAGAAGATACGGGGAACAGATTGACCTTTCTACC 384	
OY	81 HIsArGLySAStnTYrValALrGValValIsErArGAnGlnGylUsErSeRGluDeuAspLeu 100	

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:43:09 ; Search time 1038.4 Seconds  
(without alignments)  
2526.643 Million cell updates/sec

Title: US-10-082-502-21

Perfect score: 847

Sequence: 1 RNSODLHCACALVDELEW.....LCSKRTDLCDAHLRSHDEL 162

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+P2n.model -DEV=xip  
-O=/cgn2\_1/USPTO.spool/US10082502/runatc\_24032003\_135100\_6599/app.query.fasta\_1.1308  
-DB=EST -OPMT=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blowsum2 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USPR=US10082502\_@CGN\_1.1\_5425\_@runatc\_24032003\_135100\_6599 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-MARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:  
1: em\_estda:\*  
2: em\_esthum:\*  
3: em\_estln:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_liv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the distribution.  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	847	100.0	754	11	AK013568	AK013568 Mus muscu
2	847	100.0	772	11	AK007914	AK007914 Mus muscu
3	847	100.0	808	13	BI408325	BI408325 Mus muscu
4	847	100.0	843	13	BI412539	BI412539 602964955
5	847	100.0	853	11	AK013014	AK013014 Mus muscu
6	847	100.0	1255	11	AK019927	AK019927 Mus muscu
7	835	98.6	665	10	AW919569	AW919569 EST350873
8	835	98.6	674	14	BQ202181	BQ202181 UT-R-EB1-
9	835	98.6	724	14	BQ210416	BQ210416 UT-R-DY1-
10	835	98.6	802	12	BF608664	BF608664 MY1-00160
11	834	98.5	678	10	BE309953	BE309953 601091903
12	834	98.5	891	13	BI408381	BI408381 602964013
13	833	98.3	709	13	BI411532	BI411532 602965126
14	832	98.2	549	10	BE198865	BE198865 ug83411.y
15	830	98.0	588	14	BM832828	BM832828 K-EST0107
16	830	98.0	591	14	BM832888	BM832888 K-EST0107
17	830	98.0	611	13	BI712899	BI712899 LQ8F09.y
18	830	98.0	657	14	BQ211867	BQ211867 UT-R-DY1-
19	830	98.0	673	14	BQ668530	BQ668530 AGENCOURT
20	830	98.0	697	12	BF344334	BF344334 602014708
21	830	98.0	719	12	BE733900	BE733900 601568370
22	830	98.0	742	12	BG824549	BG824549 602728454
23	830	98.0	745	12	BE791763	BE791763 601581586
24	830	98.0	930	14	BQ437660	BQ437660 AGENCOURT
25	830	98.0	1117	13	BM552907	BM552907 AGENCOURT
26	825	97.4	599	13	BI791489	BI791489 LQ8F09.x
27	824	97.3	880	14	BQ421827	BQ421827 AGENCOURT
28	822	97.0	783	13	BI350436	BI350436 603192905
29	817	96.5	868	14	BQ421069	BQ421069 AGENCOURT
30	815	96.2	791	12	BG704443	BG704443 602688633
31	809	95.5	595	10	AW414039	AW414039 uc98603.x
32	809	95.5	1773	13	BI550162	BI550162 603183045
33	808	95.4	856	13	BI454180	BI454180 603170602
34	808	95.4	992	13	BI456021	BI456021 603170466
35	805	95.0	875	13	BI763950	BI763950 603049829
36	804	94.9	818	12	BE870315	BE870315 601447464
37	803	94.8	475	10	AW414257	AW414257 uc96608.y
38	803	94.8	596	10	BE134206	BE134206 uc17c09.x
39	803	94.8	887	13	BI762759	BI762759 603048045
40	801	94.6	594	10	AW908970	AW908970 uc40906.x
41	798	94.2	720	13	BI909759	BI909759 603070841
42	797	94.1	593	10	AW167673	AW167673 uc47601.x
43	794	93.7	797	13	BI762927	BI762927 603047987
44	791	93.4	746	13	BI551339	BI551339 603193144
45	790	93.3	783	13	BI763784	BI763784 603049729

## ALIGNMENTS

RESULT 1  
AK013568  
LOCUS  
DEFINITION  
AK013568 754 bp mRNA linear HTC 19-JAN-2002  
Mus musculus adult male hippocampus cDNA, RIKEN full-length  
enriched library, clone:2900022B12;transmembrane protein 4, full  
insert sequence.  
ACCESSION  
AK013568  
VERSION  
AK013568.1 GI:12850979  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus  
clone:1b;RIKEN full-length enriched mouse cDNA library  
clone:2900022B12.  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

•  
•  
•  
•

10

11

12

```
Db 165 GAATTCGCCCGGTGGACCCCAAGAACCATTCAGATGGGATCCTCCGAATCAATCCA 224
Oy 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 225 GATGGCAGCCAGATCAGTGTGTGGAGGTACCTTATGCCCGCTCAGAGGCCACCTCAGAGAG 284
Oy 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
    |||||||||||:|||||||||:|||||:|||||:|||||:|||||
Db 285 TTGCTTGGAGGAGTGTGTGACCGAATGAAGAGTAGGGGA-CAGATTGA-CCTTCTACC 342
Oy 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGluSerSerGluLeuAspLeu 100
    |||||||||||:|||||||||:|||||||||:|||||||||:|||||||||
Db 343 CACCCCAAGAACTACGTACGCGCTCGTAGACCGGAATGGAAATCCAGTGAAGTACACTTA 402
Oy 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
    |||||||||||:|||||||||:|||||||||:|||||||||:|||||||||
Db 403 CAGGCGATCCGAAATGTGACTCAGATATCANGGCGACCCCTCAAGTTGCGGTGAGAGCANT 462
Oy 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
    |||||||||||:|||||||||:|||||||||:|||||||||:|||||||||
Db 463 GTGGAAAGATACGAGATGAGCTTATCGAATCTTCTCCAGAGAGGCTGACAAAGCTTAA 522
Oy 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSer-HisAs 160
    |||||||||||:|||||||||:|||||||||:|||||||||:|||||||||
Db 523 GACAAACTTTGCGAGTAAGCGGACAGACTCTATGTGACCAATCCCTGCACAGATCTCAGCA 582
Oy 160 pGluLeu 162
    :|||||
Db 583 AGAGCTC 589
```

Search completed: March 30, 2003, 22:42:58  
Job time : 130.151 secs